GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 13:56:56 ; Search time 21 Seconds
(without alignments)
18.318 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	28	2	A27261	proteinase inhibit
2	20	100.0	37	2	G28027	protein P10 - curl
3	20	100.0	58	2	S03810	pss1 protein - Rhi
4	20	100.0	59	2	B60232	T-cell surface gly
5	20	100.0	59	2	AC3286	LSU ribosomal prot
6	20	100.0	61	2	T03065	cytochrome-c oxida
7	20	100.0	62	2	C90471	hypothetical prote
8	20	100.0	63	2	E89802	hypothetical prote
9	20	100.0	68	2	G84484	hypothetical prote
10	20	100.0	77	2	C72645	hypothetical prote
11	20	100.0	80	2	T15286	hypothetical prote
12	20	100.0	81	2	E90324	hypothetical prote
13	20	100.0	82	2	T42003	hypothetical prote
14	20	100.0	83	2	T19717	hypothetical prote
15	20	100.0	86	2	C69107	hypothetical prote
16	20	100.0	89	2	B72296	hypothetical prote
17	20	100.0	92	2	JN0349	plastoquinol-plast
18	20	100.0	93	2	B86755	prophage pi2 prote
19	20	100.0	93	2	E91217	peptidylprolyl iso
20	20	100.0	93	2	F86063	peptidylprolyl iso
21	20	100.0	93	2	S48658	peptidylprolyl iso
22	20	100.0	93	2	AE0923	peptidyl-prolyl ci
23	20	100.0	99	2	T37093	hypothetical prote
24	20	100.0	101	2	T25656	hypothetical prote
25	20	100.0	102	2	S52711	probable neuropept
26	20	100.0	103	2	E86564	YidD family [impor
27	20	100.0	103	2	H72060	yidd family - Chla
28	20	100.0	105	1	R6MXER	ribosomal protein
29	20	100.0	105	2	S04150	plastoquinol-plast
30	20	100.0	105	2	AE1848	hypothetical prote
31	20	100.0	105	2	S76770	hypothetical prote
32	20	100.0	106	2	E81606	conserved hypothet
33	20	100.0	108	2	T13133	protein gp46 - pha
34	20	100.0	108	2	T41447	very hypothetical
35	20	100.0	109	2	B70637	hypothetical prote
36	20	100.0	111	2	D72112	hypothetical prote
37	20	100.0	111	2	F86511	hypothetical prote
38	20	100.0	112	2	B69438	hypothetical prote
39	20	100.0	113	2	T30436	hypothetical prote
40	20	100.0	113	2	F90319	hypothetical prote
41	20	100.0	114	2	H87313	nitrogen regulator
42	20	100.0	115	2	F72716	hypothetical prote
43	20	100.0	115	2	G83426	hypothetical prote
44	20	100.0	116	1	QQPSHT	hypothetical prote
45	20	100.0	116	1	QQEBHT	hypothetical prote

ALIGNMENTS

RESULT 1
A27261
proteinase inhibitor 3 - sea anemone (Stichodactyla sp.) (fragments)
C;Species: Stichodactyla sp., Stoichactis sp.
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1993
C;Accession: A27261
R;Mebs, D.; Gebauer, E.
Toxicon 20, 335, 1982
A;Title: Structural studies on a proteinase inhibitor from the sea anemone Stoichactis
A;Reference number: A27261
A;Accession: A27261
A;Molecule type: protein
A;Residues: 1-28 <MEB>

Query Match 100.0%; Score 20; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 4 VLEP 7

RESULT 2
G28027
protein P10 - curled-leaved tobacco (fragment)
C;Species: Nicotiana plumbaginifolia (curled-leaved tobacco)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 11-Jan-2000
C;Accession: G28027
R;Bauw, G.; De Loose, M.; Inze, D.; Van Montagu, M.; Vandekerckhove, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 4806-4810, 1987
A;Title: Alterations in the phenotype of plant cells studied by NH2-terminal amino acid
A;Reference number: A94167
A;Accession: G28027
A;Molecule type: protein
A;Residues: 1-37 <BAU>
C;Superfamily: thaumatin I

Query Match 100.0%; Score 20; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 29 VLEP 32

RESULT 3
S03810
pss1 protein - Rhizobium leguminosarum bv. phaseoli
C;Species: Rhizobium leguminosarum bv. phaseoli
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 11-Jan-2000

C;Accession: S03810
R;Borthakur, D.; Barker, R.F.; Latchford, J.W.; Rossen, L.; Johnston, A.W.B.
Mol. Gen. Genet. 213, 155-162, 1988
A;Title: Analysis of pss genes of Rhizobium leguminosarum required for exopolysaccharide
genes.
A;Reference number: S03810; MUID:89127136; PMID:2851702
A;Accession: S03810
A;Molecule type: DNA
A;Residues: 1-58 <BOR>
A;Cross-references: EMBL:X12568; NID:g46255; PIDN:CAA31078.1; PID:g46256
C;Genetics:
A;Gene: pss1
C;Superfamily: Aquifex aeolicus cysQ protein

Query Match 100.0%; Score 20; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 54 VLEP 57

RESULT 4
B60232
T-cell surface glycoprotein CD3 delta chain precursor - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 08-Dec-1992 #sequence_revision 08-Dec-1992 #text_change 23-Jul-1999
C;Accession: B60232
R;Clevers, H.; MacHugh, N.D.; Bensaid, A.; Dunlap, S.; Baldwin, C.L.; Kaushal, A.; Iams,
Eur. J. Immunol. 20, 809-817, 1990
A;Title: Identification of a bovine surface antigen uniquely expressed on CD4-CD8- T cel
A;Reference number: A60232; MUID:90269333; PMID:1971793
A;Accession: B60232
A;Molecule type: mRNA
A;Residues: 1-59 <CLE>
A;Cross-references: GB:X53269; NID:g287745; PIDN:CAA37367.1; PID:g287746
C;Superfamily: T-cell surface glycoprotein CD3 delta chain; immunoglobulin homology
C;Keywords: glycoprotein; receptor; T-cell; transmembrane protein

Query Match 100.0%; Score 20; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 26 VLEP 29

RESULT 5
AC3286
LSU ribosomal protein L32P [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 17-May-2002
C;Accession: AC3286
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-59 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51454.1; PID:g17982164; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0272
A;Map position: I
C;Superfamily: Escherichia coli ribosomal protein L32

Query Match 100.0%; Score 20; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 54 VLEP 57

RESULT 6
T03065
cytochrome-c oxidase homolog 043L - Chilo iridescent virus
C;Species: Chilo iridescent virus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
C;Accession: T03065
R;Bahr, U.; Tidona, C.A.; Darai, G.
Virus Genes 15, 235-245, 1997
A;Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101
A;Reference number: Z14834; MUID:98141693; PMID:9482589
A;Accession: T03065
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-61 <BAH>
A;Cross-references: EMBL:AF003534; NID:g2738385; PIDN:AAB94439.1; PID:g2738412

Query Match 100.0%; Score 20; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 53 VLEP 56

RESULT 7
C90471
hypothetical protein SSO11456 [imported] - Sulfolobus solfataricus transposon ISC1058
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: C90471
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90471
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-62 <KUR>
A;Cross-references: GB:AE006641; NID:g13816289; PIDN:AAK43026.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO11456

Query Match 100.0%; Score 20; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 20 VLEP 23

RESULT 8
E89802
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: E89802
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: E89802
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <KUR>
A;Cross-references: GB:BA000018; PID:g13700278; PIDN:BAB41576.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SAS009

Query Match 100.0%; Score 20; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 35 VLEP 38

RESULT 9
G84484
hypothetical protein At2g07350 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84484
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-68 <STO>
A;Cross-references: GB:AE002093; NID:g4309758; PIDN:AAD15527.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g07350
A;Map position: 2

Query Match 100.0%; Score 20; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 6 VLEP 9

RESULT 10
C72645
hypothetical protein APES025 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: C72645
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72645
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79563.1; PID:d1043349; PID:g510
A;Experimental source: strain K1
C;Genetics:
A;Gene: APES025

Query Match 100.0%; Score 20; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 30 VLEP 33

RESULT 11
T15286
hypothetical protein M01D7.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15286
R;Gattung, S.; Goela, D.; Wilson, R.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid M01D7.
A;Reference number: Z18322
A;Accession: T15286
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-80 <GAT>
A;Cross-references: EMBL:AF003739; NID:g2105482; PID:g2105484; PIDN:AAB58066.1; GSPDB:G
A;Experimental source: strain Bristol N2; clone M01D7
C;Genetics:
A;Gene: CESP:M01D7.5
A;Map position: 1
A;Introns: 19/1; 38/3

Query Match 100.0%; Score 20; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 72 VLEP 75

RESULT 12
E90324
hypothetical protein SSO8568 [imported] - Sulfolobus solfataricus transposon ISC1229
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: E90324
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: E90324
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-81 <KUR>
A;Cross-references: GB:AE006641; NID:g13814884; PIDN:AAK41852.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO8568

Query Match 100.0%; Score 20; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 23 VLEP 26

RESULT 13
T42003
hypothetical protein H3 - human herpesvirus 7 (strain JI)
C;Species: human herpesvirus 7
A;Variety: strain JI
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C;Accession: T41904; T42003
R;Nicholas, J.
submitted to the EMBL Data Library, December 1995
A;Description: Determination and analysis of the complete nucleotide sequence of human

A;Reference number: Z22022
A;Accession: T41904
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-82 <NI2>
A;Cross-references: EMBL:U43400; PIDN:AAC54664.1
A;Experimental source: strain JI
A;Genetics: GN1
A;Accession: T42003
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-82 <NIC>
A;Cross-references: EMBL:U43400; PIDN:AAC54763.1
A;Experimental source: strain JI
A;Genetics: GN2
C;Genetics: <GN1>
A;Gene: H3
A;Map position: 3976-4224
C;Genetics: <GN2>
A;Gene: H3'
A;Map position: 143023-143271

Query Match 100.0%; Score 20; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 19 VLEP 22

RESULT 14
T19717
hypothetical protein C34D1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19717
R;Baynes, C.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19168
A;Accession: T19717
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-83 <WIL>
A;Cross-references: EMBL:Z78060; PIDN:CAB01491.1; GSPDB:GN00023; CESP:C34D1.1
A;Experimental source: clone C34D1
C;Genetics:
A;Gene: CESP:C34D1.1
A;Map position: 5
A;Introns: 35/1; 57/2

Query Match 100.0%; Score 20; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 25 VLEP 28

RESULT 15
C69107
hypothetical protein MTH1799 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: C69107
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: C69107
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-86 <MTH>
A;Cross-references: GB:AE000934; GB:AE000666; NID:g2622924; PIDN:AAB86265.1; PID:g26229
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1799

Query Match 100.0%; Score 20; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 58 VLEP 61

Search completed: November 25, 2003, 14:00:52
Job time : 23 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 13:47:36 ; Search time 11 Seconds
(without alignments)
17.101 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	59	1	CD3D_BOVIN	Q28072 bos taurus
2	20	100.0	59	1	RL32_BRUME	Q8yj14 brucella me
3	20	100.0	71	1	IF1_LEPIN	Q9xd14 leptospira
4	20	100.0	92	1	PPIC_ECOLI	P39159 escherichia
5	20	100.0	92	1	PPIC_SALTY	Q9l6s3 salmonella
6	20	100.0	95	1	YPSS_RHILP	P10497 rhizobium l
7	20	100.0	98	1	RPOZ_THEAQ	Q9evv4 thermus aqu
8	20	100.0	99	1	RPOZ_THETH	Q8rqe7 thermus the
9	20	100.0	100	1	VG6_BPB03	Q37886 bacteriopha
10	20	100.0	103	1	Y592_CHLPN	Q9z7w2 chlamydia p
11	20	100.0	105	1	RL3E_METVA	P14025 methanococc
12	20	100.0	106	1	PETD_HORVU	P12361 hordeum vul
13	20	100.0	112	1	YF07_ARCFU	O28765 archaeoglob
14	20	100.0	114	1	RLA2_EIMTE	Q967y9 eimeria ten
15	20	100.0	116	1	MERT_ACICA	Q52106 acinetobact
16	20	100.0	116	1	MERT_ALCSP	P94185 alcaligenes
17	20	100.0	116	1	MERT_PSEAE	P04140 pseudomonas
18	20	100.0	116	1	MERT_PSEFL	Q51769 pseudomonas
19	20	100.0	116	1	MERT_SALTI	P04336 salmonella
20	20	100.0	116	1	MERT_SERMA	P13112 serratia ma
21	20	100.0	116	1	REV_HV1OY	P20887 human immun
22	20	100.0	117	1	DHSD_ARCFU	O29573 archaeoglob
23	20	100.0	120	1	SYE_STAXY	P77984 staphylococ
24	20	100.0	126	1	MERT_ENTAG	P94700 enterobacte
25	20	100.0	129	1	UMP1_SCHPO	O74416 schizosacch
26	20	100.0	129	1	YIO1_YEAST	P40461 saccharomyc
27	20	100.0	134	1	WNT1_CHICK	Q91029 gallus gall
28	20	100.0	136	1	RK16_MAIZE	P08528 zea mays (m
29	20	100.0	139	1	PETD_PEA	P06527 pisum sativ
30	20	100.0	146	1	YN81_YEAST	P40342 saccharomyc
31	20	100.0	147	1	VG36_HSV11	Q00146 ictalurid h
32	20	100.0	150	1	ZFP2_ARATH	Q39261 arabidopsis
33	20	100.0	153	1	IEX1_MOUSE	P46694 mus musculu
34	20	100.0	156	1	IEX1_HUMAN	P46695 homo sapien
35	20	100.0	157	1	HES2_MOUSE	O54792 mus musculu
36	20	100.0	157	1	HES2_RAT	P35429 rattus norv
37	20	100.0	158	1	VH21_MYXVL	P28850 myxoma viru
38	20	100.0	160	1	PETD_ARATH	P56774 arabidopsis
39	20	100.0	160	1	PETD_LOTJA	Q9bbq5 lotus japon
40	20	100.0	160	1	PETD_MAIZE	P05643 zea mays (m
41	20	100.0	160	1	PETD_MARPO	P06250 marchantia
42	20	100.0	160	1	PETD_MESVI	Q9muv2 mesostigma
43	20	100.0	160	1	PETD_ORYSA	P12118 oryza sativ
44	20	100.0	160	1	PETD_SPIOL	P00166 spinacia ol
45	20	100.0	160	1	PETD_TOBAC	P06249 nicotiana t

ALIGNMENTS

RESULT 1
CD3D_BOVIN
ID CD3D_BOVIN STANDARD; PRT; 59 AA.
AC Q28072;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD3 delta chain precursor (T-cell receptor
DE T3 delta chain) (Fragment).
GN CD3D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269333; PubMed=1971793;
RA Clevers H., Machugh N.D., Bensaid A., Dunlap S., Baldwin C.L.,
RA Kaushal A., Iams K., Howard C.J., Morrison W.I.;
RT "Identification of a bovine surface antigen uniquely expressed on
RT CD4-CD8-T cell receptor gamma/delta+ T lymphocytes.";
RL Eur. J. Immunol. 20:809-817(1990).
CC -!- FUNCTION: THE CD3 COMPLEX MEDIATES SIGNAL TRANSDUCTION.
CC -!- SUBUNIT: THE TCR/CD3 COMPLEX OF T LYMPHOCYTES CONSISTS OF EITHER
CC A TCR ALPHA/BETA OR TCR GAMMA/DELTA HETERODIMER COEXPRESSED AT THE
CC CELL SURFACE WITH THE INVARIANT SUBUNITS OF CD3 LABELED GAMMA,
CC DELTA, EPSILON, ZETA, AND ETA.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC --

CC --
DR EMBL; X53269; CAA37367.1; -.
DR PIR; B60232; B60232.
KW T-cell; Receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 >59 T-CELL SURFACE GLYCOPROTEIN CD3 DELTA
FT CHAIN.
FT DOMAIN 22 >59 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 59 59
SQ SEQUENCE 59 AA; 6468 MW; DE2C500A9C99F40E CRC64;

Query Match 100.0%; Score 20; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 26 VLEP 29

```
RESULT 2
RL32_BRUME
ID   RL32_BRUME     STANDARD;      PRT;    59 AA.
AC   Q8YJI4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L32.
GN   RPMF OR BMEI0272 OR BR1775.
OS   Brucella melitensis, and
OS   Brucella suis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29459, 29461;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=B.melitensis; STRAIN=16M / ATCC 23456 / Biotype 1;
RX   MEDLINE=20020109; PubMed=11756688;
RA   DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA   Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA   Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA   Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA   Haselkorn R., Kyrpides N., Overbeek R.;
RT   "The genome sequence of the facultative intracellular pathogen
RT   Brucella melitensis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=B.suis; STRAIN=1330 / Biovar 1;
RX   MEDLINE=22247741; PubMed=12271122;
RA   Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA   Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA   Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA   Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA   Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA   Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT   "The Brucella suis genome reveals fundamental similarities between
RT   animal and plant pathogens and symbionts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC   -!- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE009470; AAL51454.1; -.
DR   EMBL; AE014468; AAN30673.1; -.
DR   PIR; AC3286; AC3286.
DR   TIGR; BR1775; -.
DR   HAMAP; MF_00340; -; 1.
DR   InterPro; IPR002677; Ribosomal_L32p.
DR   InterPro; IPR005718; S32_bact_org.
DR   Pfam; PF01783; Ribosomal_L32p; 1.
DR   TIGRFAMs; TIGR01031; rpmF_bact; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   59 AA;  6790 MW;  CE3BC2B67F7DF6DB CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         54 VLEP 57


RESULT 3
```

```
IF1_LEPIN
ID   IF1_LEPIN      STANDARD;      PRT;    71 AA.
AC   Q9XD14;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Translation initiation factor IF-1.
GN   INFA OR LA0761.
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lai / Serogroup Icterohaemorrhagiae / Serovar lai;
RX   MEDLINE=20088835; PubMed=10620683;
RA   Zuerner R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;
RT   "Characterization of the Leptospira interrogans S10-spc-alpha
RT   operon.";
RL   FEMS Microbiol. Lett. 182:303-308(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX   MEDLINE=22598143; PubMed=12712204;
RA   Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA   Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA   Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA   Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA   Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA   Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA   Xu J.-G., Zhao G.-P.;
RT   "Unique physiological and pathogenic features of Leptospira
RT   interrogans revealed by whole-genome sequencing.";
RL   Nature 422:888-893(2003).
CC   -!- FUNCTION: No specific function has so far been attributed to this
CC       initiation factor; however, it seems to stimulate more or less all
CC       the activities of the other two initiation factors, IF-2 and IF-3.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the IF-1 family.
CC   -!- SIMILARITY: Contains 1 S1-like domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF115283; AAD40605.1; -.
DR   EMBL; AE011262; AAN47960.1; -.
DR   HSSP; P02998; 1AH9.
DR   HAMAP; MF_00075; -; 1.
DR   InterPro; IPR003029; S1.
DR   InterPro; IPR006196; S1_IF1.
DR   InterPro; IPR004368; TIF_IF1.
DR   Pfam; PF00575; S1; 1.
DR   SMART; SM00316; S1; 1.
DR   TIGRFAMs; TIGR00008; infA; 1.
DR   PROSITE; PS50832; S1_IF1_TYPE; 1.
KW   Initiation factor; Protein biosynthesis; Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   DOMAIN        1     71       S1-LIKE.
SQ   SEQUENCE   71 AA;  8062 MW;  978564AB12140CD9 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 90;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         12 VLEP 15
```

```
RESULT 4
PPIC_ECOLI
ID   PPIC_ECOLI     STANDARD;      PRT;    92 AA.
AC   P39159;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Peptidyl-prolyl cis-trans isomerase C (EC 5.2.1.8) (PPIase C)
DE   (Rotamase C) (Parvulin).
GN   PPIC OR PARVA OR B3775 OR C4697 OR Z5286 OR ECS4709.
OS   Escherichia coli,
OS   Escherichia coli O6, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 217992, 83334;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE.
RX   MEDLINE=95010704; PubMed=7925971;
RA   Rahfeld J.-U., Ruecknagel K.P., Schelbert B., Ludwig B., Hacker J.,
RA   Mann K., Fischer G.;
RT   "Confirmation of the existence of a third family among
RT   peptidyl-prolyl cis/trans isomerases. Amino acid sequence and
RT   recombinant production of parvulin.";
RL   FEBS Lett. 352:180-184(1994).
RN   [2]
RP   PRELIMINARY SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=92358234; PubMed=1379743;
RA   Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT   "Analysis of the Escherichia coli genome: DNA sequence of the region
RT   from 84.5 to 86.5 minutes.";
RL   Science 257:771-778(1992).
RN   [3]
RP   REVISIONS, AND IDENTIFICATION.
RX   MEDLINE=95184297; PubMed=7878732;
RA   Rudd K.E., Sofia H.J., Koonin E.V., Plunkett G. III, Lazar S.,
RA   Rouviere P.E.;
RT   "Conserved sequence motifs in bacterial and bacteriophage
RT   chaperonins.";
RL   Trends Biochem. Sci. 20:14-15(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [7]
RP   SEQUENCE OF 1-21, AND CHARACTERIZATION.
RX   MEDLINE=94215709; PubMed=8163020;
RA   Rahfeld J.-U., Schierhorn A., Mann K., Fischer G.;
RT   "A novel peptidyl-prolyl cis/trans isomerase from Escherichia coli.";
RL   FEBS Lett. 343:65-69(1994).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS. IT PREFERS
CC       AMINO ACID RESIDUES WITH HYDROPHOBIC SIDE CHAINS LIKE LEUCINE AND
CC       PHENYLALANINE IN THE P1 POSITION OF THE PEPTIDES SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC       (omega=0).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S73874; AAB32054.1; -.
DR   EMBL; M87049; AAA67578.1; -.
DR   EMBL; AE000454; AAC76780.1; -.
DR   EMBL; AE016769; AAN83129.1; -.
DR   EMBL; AE005609; AAG58970.1; -.
DR   EMBL; AP002566; BAB38132.1; -.
DR   PIR; E91217; E91217.
DR   PIR; F86063; F86063.
DR   PIR; S48658; S48658.
DR   HSSP; Q13526; 1PIN.
DR   EcoGene; EG12352; ppiC.
DR   InterPro; IPR000297; Rotamase.
DR   Pfam; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
DR   PROSITE; PS50198; PPIC_PPIASE_2; 1.
KW   Isomerase; Rotamase; Complete proteome.
FT   INIT_MET      0      0
SQ   SEQUENCE   92 AA;  10101 MW;  AE1AE2028277DF4F CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         70 VLEP 73


RESULT 5
PPIC_SALTY
ID   PPIC_SALTY     STANDARD;      PRT;    92 AA.
AC   Q9L683;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Peptidyl-prolyl cis-trans isomerase C (EC 5.2.1.8) (PPIase C)
DE   (Rotamase C) (Parvulin).
GN   PPIC OR STM3910 OR STMD1.80 OR STY3647 OR T3388.
OS   Salmonella typhimurium, and
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602, 601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
```

```
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhi; STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS. IT PREFERS
CC       AMINO ACID RESIDUES WITH HYDROPHOBIC SIDE CHAINS LIKE LEUCINE AND
CC       PHENYLALANINE IN THE P1 POSITION OF THE PEPTIDES SUBSTRATES (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC       (omega=0).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF233324; AAF33475.1; -.
DR   EMBL; AE008882; AAL22760.1; -.
DR   EMBL; AL627279; CAD09407.1; -.
DR   EMBL; AE016845; AAO70912.1; -.
DR   HSSP; Q13526; 1PIN.
DR   StyGene; SG?????; ppiC.
DR   InterPro; IPR000297; Rotamase.
DR   Pfam; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
DR   PROSITE; PS50198; PPIC_PPIASE_2; 1.
KW   Isomerase; Rotamase; Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   92 AA;  10197 MW;  120263EFB993F924 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         70 VLEP 73


RESULT 6
YPSS_RHILP
ID   YPSS_RHILP     STANDARD;      PRT;    95 AA.
AC   P10497;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Hypothetical protein in pss 5'region (Fragment).
OS   Rhizobium leguminosarum (biovar phaseoli).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=385;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8002;
RX   MEDLINE=89127136; PubMed=2851702;
RA   Borthakur D., Barker R.F., Latchford J.W., Rossen L., Johnston A.W.B.;
RT   "Analysis of pss genes of Rhizobium leguminosarum required for
RT   exopolysaccharide synthesis and nodulation of peas: their primary
RT   structure and their interaction with psi and other nodulation
RT   genes.";
RL   Mol. Gen. Genet. 213:155-162(1988).
CC   -!- SIMILARITY: Belongs to the inositol monophosphatase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X12568; CAA31078.1; ALT_INIT.
DR   InterPro; IPR000760; Inositol_P.
DR   Pfam; PF00459; inositol_P; 1.
DR   ProDom; PD023420; Inositol_P; 1.
DR   PROSITE; PS00629; IMP_1; PARTIAL.
DR   PROSITE; PS00630; IMP_2; 1.
KW   Hypothetical protein; Exopolysaccharide synthesis; Nodulation.
FT   NON_TER       1      1
SQ   SEQUENCE   95 AA;  10149 MW;  A3DED804E1BF924B CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         91 VLEP 94


RESULT 7
RPOZ_THEAQ
ID   RPOZ_THEAQ     STANDARD;      PRT;    98 AA.
AC   Q9EVV4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   DNA-directed RNA polymerase omega chain (EC 2.7.7.6) (Transcriptase
DE   omega chain) (RNA polymerase omega subunit).
GN   RPOZ.
OS   Thermus aquaticus.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=271;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20566669; PubMed=11114902;
RA   Minakhin L., Nechaev S., Campbell E.A., Severinov K.;
RT   "Recombinant Thermus aquaticus RNA polymerase, a new tool for
RT   structure-based analysis of transcription.";
RL   J. Bacteriol. 183:71-76(2001).
RN   [2]
RP   SEQUENCE FROM N.A., SEQUENCE OF 1-10, FUNCTION, AND 3D-STRUCTURE
RP   MODELLING.
RX   MEDLINE=21107642; PubMed=11158566;
```

```
RA   Minakhin L., Bhagat S., Brunning A., Campbell E.A., Darst S.A.,
RA   Ebright R.H., Severinov K.;
RT   "Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase
RT   subunit RPB6 are sequence, structural, and functional homologs and
RT   promote RNA polymerase assembly.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:892-897(2001).
CC   -!- FUNCTION: Promotes RNA polymerase assembly. Latches the N- and C-
CC       terminal regions of the beta' subunit thereby facilitating its
CC       interaction with the beta and alpha subunits.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: Consists of a sigma factor and the RNAP core enzyme which
CC       is composed of 2 alpha chains, 1 beta chain, 1 beta' chain and 1
CC       omega chain.
CC   -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE OMEGA CHAIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ295839; CAC15848.1; -.
DR   PDB; 1HQM; 07-FEB-01.
DR   PDB; 1L9Z; 31-MAY-02.
DR   HAMAP; MF_00366; -; 1.
DR   InterPro; IPR003716; RNA_pol_omega.
DR   InterPro; IPR006110; RNA_poly_Rpb6.
DR   Pfam; PF01192; RNA_pol_Rpb6; 1.
DR   TIGRFAMs; TIGR00690; rpoZ; 1.
KW   Transferase; DNA-directed RNA polymerase; Transcription;
KW   3D-structure.
FT   INIT_MET      0      0
FT   TURN          3      4
FT   HELIX         5     11
FT   HELIX        15     26
FT   TURN         27     28
FT   HELIX        29     31
FT   TURN         32     32
FT   TURN         35     36
FT   TURN         40     42
FT   STRAND       46     46
FT   TURN         50     51
FT   STRAND       55     55
FT   HELIX        59     67
FT   TURN         68     69
FT   HELIX        81     86
FT   TURN         87     90
SQ   SEQUENCE   98 AA;  11494 MW;  8E4529F14BB0C266 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         38 VLEP 41


RESULT 8
RPOZ_THETH
ID   RPOZ_THETH     STANDARD;      PRT;    99 AA.
AC   Q8RQE7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable DNA-directed RNA polymerase omega chain (EC 2.7.7.6)
DE   (Transcriptase omega chain) (RNA polymerase omega subunit).
GN   RPOZ.
OS   Thermus thermophilus.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=274;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB8 / ATCC 27634;
RA   Vassylyeva M.N., Lee J., Sekine S., Laptenko O., Kuramitsu S.,
RA   Shibata T., Inoue Y., Borukov S., Vassylyev D.G., Yokoyama S.;
RT   "Purification, crystallization and initial crystallographic analysis
RT   of RNA polymerase holoenzyme from Thermus thermophilus.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Promotes RNA polymerase assembly. Latches the N- and C-
CC       terminal regions of the beta' subunit thereby facilitating its
CC       interaction with the beta and alpha subunits (By similarity).
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: Consists of a sigma factor and the RNAP core enzyme which
CC       is composed of 2 alpha chains, 1 beta chain, 1 beta' chain and 1
CC       omega chain (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE OMEGA CHAIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB083791; BAB89402.1; -.
DR   HAMAP; MF_00366; -; 1.
DR   InterPro; IPR003716; RNA_pol_omega.
DR   InterPro; IPR006110; RNA_poly_Rpb6.
DR   Pfam; PF01192; RNA_pol_Rpb6; 1.
DR   TIGRFAMs; TIGR00690; rpoZ; 1.
KW   Transferase; DNA-directed RNA polymerase; Transcription.
SQ   SEQUENCE   99 AA;  11505 MW;  824E3933AA09912E CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         39 VLEP 42


RESULT 9
VG6_BPB03
ID   VG6_BPB03      STANDARD;      PRT;   100 AA.
AC   Q37886;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Early protein GP6.
GN   6.
OS   Bacteriophage B103.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC   phi-29-like viruses.
OX   NCBI_TaxID=10778;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98019084; PubMed=9358052;
RA   Pecenkova T., Benes V., Paces J., Vlcek C., Paces V.;
RT   "Bacteriophage B103: complete DNA sequence of its genome and
RT   relationship to other Bacillus phages.";
RL   Gene 199:157-163(1997).
CC   -!- FUNCTION: ESSENTIAL FOR VIRAL DNA REPLICATION. ACTIVATES THE
CC       INITIATION OF DNA REPLICATION BY FORMING A MULTIMERIC
CC       NUCLEOPROTEIN COMPLEX AT THE REPLICATION ORIGINS (BY SIMILARITY).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X99260; CAA67653.1; -.
KW   Early protein; DNA replication.
SQ   SEQUENCE   100 AA;  11367 MW;  DFA32866644580E8 CRC64;

  Query Match            100.0%;  Score 20;  DB 1;  Length 100;
  Best Local Similarity  100.0%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         30 VLEP 33


RESULT 10
Y592_CHLPN
ID   Y592_CHLPN     STANDARD;      PRT;   103 AA.
AC   Q9Z7W2; Q9K2D2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Hypothetical UPF0161 protein CPn0592/CP0156/CPj0592.
GN   CPN0592 OR CP0156 OR CPJ0592.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SIMILARITY: Belongs to the UPF0161 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001643; AAD18731.1; -.
DR   EMBL; AE002177; AAF38037.1; ALT_INIT.
DR   EMBL; AP002547; BAA98799.1; -.
DR   PIR; E86564; E86564.
DR   PIR; H72060; H72060.
DR   TIGR; CP0156; -.
DR   HAMAP; MF_00386; -; 1.
DR   InterPro; IPR002696; DUF37.
DR   Pfam; PF01809; DUF37; 1.
DR   ProDom; PD004225; DUF37; 1.
DR   TIGRFAMs; TIGR00278; TIGR00278; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   103 AA;  11751 MW;  CE9CA4852EA15A7C CRC64;

  Query Match            100.0%;  Score 20;  DB 1;  Length 103;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         86 VLEP 89


RESULT 11
RL3E_METVA
ID   RL3E_METVA     STANDARD;      PRT;   105 AA.
AC   P14025;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L30e.
GN   RPL30E.
OS   Methanococcus vannielii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanococcaceae; Methanococcus.
OX   NCBI_TaxID=2187;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 1224;
RX   MEDLINE=89362493; PubMed=2475640;
RA   Lechner K., Heller G., Boeck A.;
RT   "Organization and nucleotide sequence of a transcriptional unit of
RT   Methanococcus vannielii comprising genes for protein synthesis
RT   elongation factors and ribosomal proteins.";
RL   J. Mol. Evol. 29:20-27(1989).
CC   -!- SIMILARITY: BELONGS TO THE L30E FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15970; CAA34087.1; -.
DR   PIR; S06621; R6MXER.
DR   HSSP; P14120; 1CN9.
DR   HAMAP; MF_00481; -; 1.
DR   InterPro; IPR000231; Ribosomal_L30e.
DR   InterPro; IPR004038; Ribosomal_L7A.
DR   Pfam; PF01248; Ribosomal_L7Ae; 1.
DR   ProDom; PD004495; Ribosomal_L30e; 1.
DR   PROSITE; PS00709; RIBOSOMAL_L30E_1; 1.
DR   PROSITE; PS00993; RIBOSOMAL_L30E_2; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   105 AA;  11515 MW;  F1CAF5380E358EC8 CRC64;

  Query Match            100.0%;  Score 20;  DB 1;  Length 105;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         91 VLEP 94
```

```
RESULT 12
PETD_HORVU
ID   PETD_HORVU     STANDARD;      PRT;   106 AA.
AC   P12361;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cytochrome B6-F complex subunit 4 (17 kDa polypeptide) (Fragment).
GN   PETD.
OS   Hordeum vulgare (Barley).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC   STRAIN=cv. Sabarlis;
RX   MEDLINE=89240047; PubMed=2654887;
RA   Andreeva A.V., Buryakova A.A., Reverdatto S.V., Chakhmakhcheva O.G.,
RA   Efimov V.A.;
RT   "Nucleotide sequence of the 5.2 kbp barley chloroplast DNA fragment,
RT   containing psbB-psbH-petB-petD gene cluster.";
RL   Nucleic Acids Res. 17:2859-2860(1989).
CC   -!- FUNCTION: THIS POLYPEPTIDE OF UNKNOWN FUNCTION IS ONE OF THE
CC       COMPONENTS OF THE CYTOCHROME B6-F COMPLEX.
CC   -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6, 17
CC       kDa POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P12361-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P12361-2; Sequence=VSP_001239;
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY. CORRESPONDS TO THE
CC       CARBOXYL END OF MITOCHONDRIAL CYTOCHROME B.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X14107; CAA32268.1; -.
DR   EMBL; X14107; CAA32269.1; -.
DR   PIR; JN0349; JN0349.
DR   InterPro; IPR005870; Cytb6/f_IV.
DR   InterPro; IPR005798; Cytb_b6_C.
DR   Pfam; PF00032; cytochrome_b_C; 1.
DR   TIGRFAMs; TIGR01156; cytb6/f_IV; 1.
DR   PROSITE; PS00193; CYTOCHROME_B_QO; 1.
KW   Electron transport; Chloroplast; Photosynthesis; Transmembrane;
KW   Alternative splicing.
FT   VARSPLIC      1     17       MSGSFGGWILKSSPIPI -> MGV (in isoform
FT                                Short).
FT                                /FTId=VSP_001239.
FT   NON_TER     106    106
SQ   SEQUENCE   106 AA;  11671 MW;  D4D83AE9C47DFD3D CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 106;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         70 VLEP 73


RESULT 13
YF07_ARCFU
ID   YF07_ARCFU     STANDARD;      PRT;   112 AA.
AC   O28765;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AF1507.
GN   AF1507.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000998; AAB89743.1; -.
DR   PIR; B69438; B69438.
DR   TIGR; AF1507; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   112 AA;  13009 MW;  6C820B43E141B4DB CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         82 VLEP 85


RESULT 14
RLA2_EIMTE
ID   RLA2_EIMTE     STANDARD;      PRT;   114 AA.
AC   Q967Y9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   60S acidic ribosomal protein P2.
OS   Eimeria tenella.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC   Eimeria.
OX   NCBI_TaxID=5802;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAPt38;
RA   Labbe M., Pery P.;
RT   "Molecular cloning of a cDNA encoding an acidic ribosomal protein P2
RT   of Eimeria tenella.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC       PROTEIN SYNTHESIS (By similarity).
```

```
CC   -!- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC       SUBUNIT (By similarity).
CC   -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF353514; AAK38885.1; ALT_INIT.
DR   InterPro; IPR001813; 60s_ribosomal.
DR   Pfam; PF00428; 60s_ribosomal; 1.
KW   Ribosomal protein; Phosphorylation.
SQ   SEQUENCE   114 AA;  11444 MW;  4C08C3C569078AA9 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 114;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         26 VLEP 29


RESULT 15
MERT_ACICA
ID   MERT_ACICA     STANDARD;      PRT;   116 AA.
AC   Q52106;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mercuric transport protein (Mercury ion transport protein).
GN   MERT.
OS   Acinetobacter calcoaceticus.
OG   Plasmid pKLH2.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Moraxellaceae; Acinetobacter.
OX   NCBI_TaxID=471;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94134837; PubMed=8302940;
RA   Kholodii G.Y., Lomovskaya O.L., Gorlenko Z.M., Mindlin S.Z.,
RA   Yurieva O.V., Nikiforov V.G.;
RT   "Molecular characterization of an aberrant mercury resistance
RT   transposable element from an environmental Acinetobacter strain.";
RL   Plasmid 30:303-308(1993).
CC   -!- FUNCTION: INVOLVED IN MERCURIC TRANSPORT. PASSES A HG(2+) ION
CC       FROM THE PERIPLASMIC MERP PROTEIN TO THE MERCURIC REDUCTASE
CC       (MERA).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF213017; AAA19679.1; -.
DR   InterPro; IPR003457; Transprt_MerT.
DR   Pfam; PF02411; MerT; 1.
KW   Transport; Mercuric resistance; Inner membrane; Mercury; Plasmid;
KW   Transmembrane.
FT   TRANSMEM     16     36       POTENTIAL.
FT   TRANSMEM     46     66       POTENTIAL.
FT   TRANSMEM     94    114       POTENTIAL.
FT   METAL        24     24       HG(2+) (BY SIMILARITY).
FT   METAL        25     25       HG(2+) (BY SIMILARITY).
FT   METAL        76     76       HG(2+) (BY SIMILARITY).
FT   METAL        82     82       HG(2+) (BY SIMILARITY).
SQ   SEQUENCE   116 AA;  12510 MW;  2930A92CF88EB10F CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         46 VLEP 49


Search completed: November 25, 2003, 13:59:29
Job time : 11 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 13:56:11 ; Search time 35 Seconds
(without alignments)
29.492 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	25	5	Q9BM55	Q9bm55 chione canc
2	20	100.0	26	5	Q9BM52	Q9bm52 moniliformi
3	20	100.0	30	13	Q9W7I9	Q9w7i9 gallus gall
4	20	100.0	43	12	Q9WNJ6	Q9wnj6 canine dist
5	20	100.0	47	8	Q95F85	Q95f85 spartina ma
6	20	100.0	47	8	Q95F87	Q95f87 spartina al
7	20	100.0	47	8	Q95F86	Q95f86 spartina an
8	20	100.0	52	2	Q9KWW3	Q9kww3 pseudomonas
9	20	100.0	54	16	Q8KES4	Q8kes4 chlorobium
10	20	100.0	55	4	Q96FP4	Q96fp4 homo sapien
11	20	100.0	58	12	O39631	O39631 calicivirus
12	20	100.0	58	12	O39632	O39632 calicivirus
13	20	100.0	61	12	O55728	O55728 chilo iride
14	20	100.0	62	17	Q97US6	Q97us6 sulfolobus
15	20	100.0	63	10	Q9FEX8	Q9fex8 hordeum vul
16	20	100.0	63	16	Q99WL3	Q99wl3 staphylococ
17	20	100.0	63	16	Q8NY88	Q8ny88 staphylococ
18	20	100.0	63	16	Q8CQP3	Q8cqp3 staphylococ
19	20	100.0	64	2	O05120	O05120 methylobact
20	20	100.0	65	3	O14380	O14380 schizosacch
21	20	100.0	66	16	Q8G062	Q8g062 brucella su
22	20	100.0	66	16	Q8F6P0	Q8f6p0 leptospira
23	20	100.0	68	10	Q9ZQK7	Q9zqk7 arabidopsis
24	20	100.0	71	13	Q90WY7	Q90wy7 coturnix co
25	20	100.0	71	16	Q8K619	Q8k619 streptococc
26	20	100.0	73	5	Q8IF61	Q8if61 trypanosoma
27	20	100.0	74	11	Q9JJ82	Q9jj82 mus musculu
28	20	100.0	76	4	Q8NFR3	Q8nfr3 homo sapien
29	20	100.0	77	2	Q8GMU1	Q8gmu1 acinetobact
30	20	100.0	77	16	Q99QQ0	Q99qq0 streptomyce
31	20	100.0	77	17	Q9YEI3	Q9yei3 aeropyrum p
32	20	100.0	79	11	P97585	P97585 rattus norv
33	20	100.0	80	5	O01970	O01970 caenorhabdi
34	20	100.0	80	16	Q8P0P9	Q8p0p9 streptococc
35	20	100.0	80	16	Q8FNH6	Q8fnh6 corynebacte
36	20	100.0	81	2	Q8L176	Q8l176 rhodococcus
37	20	100.0	81	12	Q9QQ86	Q9qq86 canine dist
38	20	100.0	81	17	Q97XS0	Q97xs0 sulfolobus
39	20	100.0	82	8	Q33299	Q33299 zea mays (m
40	20	100.0	82	12	Q89530	Q89530 human herpe
41	20	100.0	83	5	Q18436	Q18436 caenorhabdi
42	20	100.0	83	10	Q8LHJ5	Q8lhj5 oryza sativ
43	20	100.0	83	16	Q8VWD5	Q8vwd5 streptomyce
44	20	100.0	84	12	Q8V4C4	Q8v4c4 hepatitis c
45	20	100.0	84	17	Q8J2X4	Q8j2x4 pyrococcus

ALIGNMENTS

RESULT 1
Q9BM55
ID Q9BM55 PRELIMINARY; PRT; 25 AA.
AC Q9BM55;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE LINE-like reverse transcriptase (Fragment).
OS Chione cancellata.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida;
OC Veneroidea; Veneridae; Chione.
OX NCBI_TaxID=145464;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=LRT-L1 retrotransposon;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and ancient asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY013938; AAG59923.1; -.
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2792 MW; 119E15FE9C4A62A7 CRC64;

Query Match 100.0%; Score 20; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VLEP 4
              ||||
Db          2 VLEP 5
```

RESULT 2
Q9BM52
ID Q9BM52 PRELIMINARY; PRT; 26 AA.
AC Q9BM52;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LINE-like reverse transcriptase (Fragment).
OS   Moniliformis moniliformis.
OC   Eukaryota; Metazoa; Acanthocephala; Archiacanthocephala;
OC   Moniliformida; Moniliformidae; Moniliformis.
OX   NCBI_TaxID=10237;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TRANSPOSON=LRT-L1 retrotransposon;
RX   MEDLINE=20570504; PubMed=11121049;
RA   Arkhipova I., Meselson M.;
RT   "Transposable elements in sexual and ancient asexual taxa.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR   EMBL; AY013942; AAG59926.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      26     26
SQ   SEQUENCE   26 AA;  2757 MW;  438F0469FFA3CFAD CRC64;

  Query Match             100.0%;  Score 20;  DB 5;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         15 VLEP 18


RESULT 3
Q9W7I9
ID   Q9W7I9      PRELIMINARY;      PRT;    30 AA.
AC   Q9W7I9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   Polymerase (Fragment).
GN   POL.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99292880; PubMed=10364336;
RA   Tsang S.X., Switzer W.M., Shanmugam V., Johnson J.A., Goldsmith C.,
RA   Wright A., Fadly A., Thea D., Jaffe H., Folks T.M., Heneine W.;
RT   "Evidence of avian leukosis virus subgroup E and endogenous avian
RT   virus in measles and mumps vaccines derived from chicken cells:
RT   investigation of transmission to vaccine recipients.";
RL   J. Virol. 73:5843-5851(1999).
DR   EMBL; AF087830; AAD37027.1; -.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3401 MW;  6D3499EDD72FAC7A CRC64;

  Query Match             100.0%;  Score 20;  DB 13;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         15 VLEP 18


RESULT 4
Q9WNJ6
ID   Q9WNJ6      PRELIMINARY;      PRT;    43 AA.
AC   Q9WNJ6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Fusion protein (Fragment).
GN   F.
OS   Canine distemper virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11232;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fickel J., Czupalla O.;
RT   "Canine distemper virus in a badger (Meles meles).";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF152906; AAD38054.2; -.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
FT   NON_TER       1      1
FT   NON_TER      43     43
SQ   SEQUENCE   43 AA;  4784 MW;  DA940A521273A8C5 CRC64;

  Query Match             100.0%;  Score 20;  DB 12;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 4.6e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         30 VLEP 33


RESULT 5
Q95F85
ID   Q95F85      PRELIMINARY;      PRT;    47 AA.
AC   Q95F85;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ribosomal protein L16 (Fragment).
GN   RPL16.
OS   Spartina maritima.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Chloridoideae; Cynodonteae; Spartina.
OX   NCBI_TaxID=49786;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Baumel A., Ainouche M.L., Levasseur J.-E.;
RT   "Molecular investigations in populations of Spartina anglica C.E.
RT   Hubbard invading coastal Brittany (France).";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF276970; AAK71926.1; -.
DR   InterPro; IPR000114; Ribosomal_L16.
DR   Pfam; PF00252; Ribosomal_L16; 1.
DR   PRINTS; PR00060; RIBOSOMALL16.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5679 MW;  DD8E3D9A67537721 CRC64;

  Query Match             100.0%;  Score 20;  DB 8;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         35 VLEP 38


RESULT 6
Q95F87
ID   Q95F87      PRELIMINARY;      PRT;    47 AA.
AC   Q95F87;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Ribosomal protein L16 (Fragment).
```

```
GN   RPL16.
OS   Spartina alterniflora (smooth cordgrass).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Chloridoideae; Cynodonteae; Spartina.
OX   NCBI_TaxID=29706;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Baumel A., Ainouche M.L., Levasseur J.-E.;
RT   "Molecular investigations in populations of Spartina anglica C.E.
RT   Hubbard invading coastal Brittany (France).";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF276968; AAK71924.1; -.
DR   InterPro; IPR000114; Ribosomal_L16.
DR   Pfam; PF00252; Ribosomal_L16; 1.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5679 MW;  DD8E3D9A67537721 CRC64;

  Query Match           100.0%;  Score 20;  DB 8;  Length 47;
  Best Local Similarity  100.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         35 VLEP 38


RESULT 7
Q95F86
ID   Q95F86      PRELIMINARY;      PRT;    47 AA.
AC   Q95F86;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Ribosomal protein L16 (Fragment).
GN   RPL16.
OS   Spartina anglica.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Chloridoideae; Cynodonteae; Spartina.
OX   NCBI_TaxID=49785;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Baumel A., Ainouche M.L., Levasseur J.-E.;
RT   "Molecular investigations in populations of Spartina anglica C.E.
RT   Hubbard invading coastal Brittany (France).";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF276969; AAK71925.1; -.
DR   InterPro; IPR000114; Ribosomal_L16.
DR   Pfam; PF00252; Ribosomal_L16; 1.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5679 MW;  DD8E3D9A67537721 CRC64;

  Query Match           100.0%;  Score 20;  DB 8;  Length 47;
  Best Local Similarity  100.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         35 VLEP 38


RESULT 8
Q9KWW3
ID   Q9KWW3      PRELIMINARY;      PRT;    52 AA.
AC   Q9KWW3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Mercuric ion transport protein (Fragment).
GN   MERT.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=303;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MU10-2; TRANSPOSON=Tn5041D;
RA   Kholodii G.Y., Mindlin S.Z., Gorlenko Z.M., Bass I.A., Kalyaeva E.S.,
RA   Nikiforov V.;
RT   "Host-dependent transposition of Tn5041.";
RL   Russ. J. Genet. 36:365-373(2000).
DR   EMBL; Y18977; CAB81562.1; -.
DR   InterPro; IPR003457; Transprt_MerT.
DR   Pfam; PF02411; MerT; 1.
FT   NON_TER      52     52
SQ   SEQUENCE   52 AA;  5209 MW;  E6700E9239CBB91D CRC64;

  Query Match           100.0%;  Score 20;  DB 2;  Length 52;
  Best Local Similarity  100.0%;  Pred. No. 5.6e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         46 VLEP 49


RESULT 9
Q8KES4
ID   Q8KES4      PRELIMINARY;      PRT;    54 AA.
AC   Q8KES4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein CT0608.
GN   CT0608.
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR   EMBL; AE012834; AAM71850.1; -.
DR   TIGR; CT0608; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   54 AA;  6056 MW;  79E3572072538A47 CRC64;

  Query Match           100.0%;  Score 20;  DB 16;  Length 54;
  Best Local Similarity  100.0%;  Pred. No. 5.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         42 VLEP 45


RESULT 10
Q96FP4
```

```
ID   Q96FP4      PRELIMINARY;      PRT;    55 AA.
AC   Q96FP4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010607; AAH10607.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   55 AA;  6011 MW;  191A77612A78B7EE CRC64;

  Query Match             100.0%;  Score 20;  DB 4;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 6e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          6 VLEP 9


RESULT 11
O39631
ID   O39631      PRELIMINARY;      PRT;    58 AA.
AC   O39631;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Capsid protein (Fragment).
OS   Calicivirus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.
OX   NCBI_TaxID=11975;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B453/92/UK;
RX   MEDLINE=97278118; PubMed=9131452;
RA   Green S.M., Lambden P.R., Owen Caul E., Clarke I.N.;
RT   "Capsid sequence diversity in small round structured viruses from
RT   recent UK outbreaks of gastroenteritis.";
RL   J. Med. Virol. 52:14-19(1997).
DR   EMBL; Z73992; CAA98298.1; -.
DR   InterPro; IPR004005; Calici_coat.
DR   Pfam; PF00915; Calici_coat; 1.
FT   NON_TER       1      1
FT   NON_TER      58     58
SQ   SEQUENCE   58 AA;  6086 MW;  CA987C393B9CC02D CRC64;

  Query Match             100.0%;  Score 20;  DB 12;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         50 VLEP 53


RESULT 12
O39632
ID   O39632      PRELIMINARY;      PRT;    58 AA.
AC   O39632;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Capsid protein (Fragment).
OS   Calicivirus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.
OX   NCBI_TaxID=11975;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Carmarthen/94/UK;
RX   MEDLINE=97278118; PubMed=9131452;
RA   Green S.M., Lambden P.R., Owen Caul E., Clarke I.N.;
RT   "Capsid sequence diversity in small round structured viruses from
RT   recent UK outbreaks of gastroenteritis.";
RL   J. Med. Virol. 52:14-19(1997).
DR   EMBL; Z73997; CAA98303.1; -.
DR   InterPro; IPR004005; Calici_coat.
DR   Pfam; PF00915; Calici_coat; 1.
FT   NON_TER       1      1
FT   NON_TER      58     58
SQ   SEQUENCE   58 AA;  6207 MW;  A39F892FC412463D CRC64;

  Query Match             100.0%;  Score 20;  DB 12;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         50 VLEP 53


RESULT 13
O55728
ID   O55728      PRELIMINARY;      PRT;    61 AA.
AC   O55728;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Hypothetical 7.4 kDa protein.
OS   Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC   Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX   NCBI_TaxID=10488;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bahr U., Tidona C.A., Darai G.;
RL   Virus Genes 0:0-0(1997).
DR   EMBL; AF303741; AAB94439.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   61 AA;  7385 MW;  8F718A758C14AE0A CRC64;

  Query Match             100.0%;  Score 20;  DB 12;  Length 61;
  Best Local Similarity   100.0%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         53 VLEP 56


RESULT 14
Q97US6
ID   Q97US6      PRELIMINARY;      PRT;    62 AA.
AC   Q97US6;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Second ORF in transposon ISC1058.
GN   SSO11456.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
```

```
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR   EMBL; AE006883; AAK43026.1; -.
KW   Complete proteome.
SQ   SEQUENCE   62 AA;  6561 MW;  62BCBE0BB13F43FF CRC64;

  Query Match             100.0%;  Score 20;  DB 17;  Length 62;
  Best Local Similarity   100.0%;  Pred. No. 6.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         20 VLEP 23


RESULT 15
Q9FEX8
ID   Q9FEX8      PRELIMINARY;      PRT;    63 AA.
AC   Q9FEX8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Putative lectin (Fragment).
GN   HL#2.
OS   Hordeum vulgare (Barley).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Etiolated coleoptile;
RA   Kloppstech K.R.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Etiolated coleoptile;
RA   Grunwald I.;
RL   Thesis (2001), Department of Biological Sciences,
RL   University of Hannover, Hannover, Germany.
DR   EMBL; AJ303112; CAC19669.1; -.
KW   Lectin.
FT   NON_TER       1      1
FT   NON_TER      63     63
SQ   SEQUENCE   63 AA;  6948 MW;  D794D32B1A5A8C93 CRC64;

  Query Match             100.0%;  Score 20;  DB 10;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         54 VLEP 57


Search completed: November 25, 2003, 14:00:18
Job time : 36 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 13:46:45 ; Search time 42 Seconds
(without alignments)
15.117 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	4	21	AAB28663	Chemotactic peptid
2	20	100.0	4	22	AAB29558	Adhesion-modulator
3	20	100.0	4	24	ABP58757	Macrophage recruit
4	20	100.0	6	21	AAB28661	Chemotactic peptid
5	20	100.0	7	24	ABP58014	Prostate cancer ma
6	20	100.0	8	17	AAW06332	N-terminus of rece
7	20	100.0	8	17	AAR94654	Fragment of recept
8	20	100.0	8	24	ABP99157	ErbB2 cell overexp
9	20	100.0	9	16	AAR84629	Wheat acetyl-Coenz
10	20	100.0	10	20	AAY07169	HLA binding peptid
11	20	100.0	10	20	AAY07127	HLA binding peptid
12	20	100.0	10	24	ABP71117	Ced-3 protein CARD
13	20	100.0	14	22	AAM00583	Human transcriptio
14	20	100.0	17	21	AAB08085	Amino acid sequenc
15	20	100.0	17	21	AAB08086	Amino acid sequenc
16	20	100.0	17	24	ABP56417	E. coli valS relat
17	20	100.0	19	19	AAW70175	Internal peptide f
18	20	100.0	19	22	AAE05631	Internal peptide f
19	20	100.0	20	22	AAB50346	Adeno-associated v
20	20	100.0	25	16	AAR79207	TGF-beta5 residues
21	20	100.0	25	19	AAW67273	Peptide #20 with c
22	20	100.0	26	15	AAR52310	Mouse heavy chain
23	20	100.0	27	22	ABG51099	Human liver peptid
24	20	100.0	27	22	ABB31072	Peptide #3723 enco
25	20	100.0	27	22	ABB36261	Peptide #3767 enco
26	20	100.0	27	22	ABB21633	Protein #3632 enco
27	20	100.0	27	22	AAM57035	Human brain expres
28	20	100.0	27	22	AAM69425	Human bone marrow
29	20	100.0	27	22	AAM17259	Peptide #3693 enco
30	20	100.0	27	22	AAM29756	Peptide #3793 enco
31	20	100.0	27	22	AAM04951	Peptide #3633 enco
32	20	100.0	27	23	ABG39043	Human peptide enco
33	20	100.0	30	18	AAW24711	Finger 2 domain of
34	20	100.0	34	16	AAR66400	GalNAc-transferase
35	20	100.0	34	18	AAW16483	N-acetylgalactosam
36	20	100.0	34	21	AAB02780	Xenopus laevis TGF
37	20	100.0	34	21	AAY92551	Finger 2 subdomain
38	20	100.0	34	22	AAM96570	Human reproductive
39	20	100.0	35	21	AAY89142	Core polypeptide f
40	20	100.0	35	22	ABB00501	Viral DP178/107-li
41	20	100.0	35	22	ABB01977	Viral core polypep
42	20	100.0	35	22	AAU13050	DP178-like/DP107-l
43	20	100.0	35	22	AAB77497	Core polypeptide T
44	20	100.0	36	21	AAB09513	Xenopus TGF-beta 5
45	20	100.0	37	22	ABG58019	Human liver peptid

ALIGNMENTS

RESULT 1
AAB28663
ID AAB28663 standard; peptide; 4 AA.
XX
AC AAB28663;
XX
DT 13-FEB-2001 (first entry)
XX
DE Chemotactic peptide pepJ'.
XX
KW Chemotactic; osteopontin; vulnerary; antiarthritic; antipsoriatic;
KW cytostatic; antitumour; antiinflammatory; osteopathic;
KW wound healing; cell migration; chemotaxis; atherosclerosis; cancer;
KW angiogenic-associated disease; arthritis; psoriasis; haemangioma;
KW ocular neovascularisation; cell apoptosis; nitrous oxide production;
KW inflammation; osteoporosis; immune disease.
XX
OS Mammalia.
OS Synthetic.
XX
PN WO200063247-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-US10344.
XX
PR 15-APR-1999; 99US-0129764.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX

DR WPI; 2000-687159/67.
XX
PT New osteopontin-derived chemotactic and inhibitory peptides, useful for
PT promoting scarless wound healing, modulating cellular chemotaxis,
PT treating formation of atherosclerotic plaques and preventing metastasis
PT -
XX
PS Claim 17; Page 43; 54pp; English.
XX
CC The present sequence is an osteopontin-derived chemotactic peptide.
CC Such chemotactic peptides are useful for promoting scarless wound
CC healing, modulating chemotaxis and promoting cell migration to a target
CC site in a cell of a subject. They are also used for modulating cellular
CC chemotaxis in a mammalian cell such as smooth muscle cell, a macrophage,
CC an endothelial cell, a vascular cell and a tumourigenic cell. They are
CC useful for treating the formation of atherosclerotic plaques in a
CC subject. The peptides are used for preventing metastasis, treating an
CC angiogenic-associated disease such as arthritis, psoriasis, haemangioma,
CC tumour metastasis or ocular neovascularisation. They are also used for
CC activating cell apoptosis, for modulating nitrous oxide production and
CC for inducing chemotaxis. The peptides are useful for diagnosing, treating
CC and preventing tumour metastasis, inflammation, osteoporosis and immune
CC diseases. They can also be used to enhance an immune response by
CC attracting macrophages.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 2
AAB29558
ID AAB29558 standard; peptide; 4 AA.
XX
AC AAB29558;
XX
DT 14-FEB-2001 (first entry)
XX
DE Adhesion-modulatory peptide, SEQ ID NO:15.
XX
KW Adhesion modulatory peptide; target cell adhesion; cell adhesion;
KW endothelial cell; fibroblast; macrophage; neutrophil; myofibroblast;
KW collagen; glycosaminoglycan; extracellular matrix; synthetic substrate;
KW vascular growth; wound healing; keloid formation; scarring; fibrosis;
KW anticoagulant; vulnerary; immunomodulatory; antibacterial; anticancer;
KW antitumorigenic; anti-CD44 activity; prosthesis; implant.
XX
OS Synthetic.
XX
PN WO200063236-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-US10329.
XX
PR 16-APR-1999; 99US-0129709.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX
DR WPI; 2001-007002/01.
XX
PT Novel adhesive modulatory peptides useful for modulating adhesion of
PT target cells such as endothelial cells, fibroblasts, macrophages to
PT substrate such as polyvinyl surfaces, collagen -
XX
PS Claim 8; Page 27; 35pp; English.
XX
CC The invention relates to novel adhesion modulatory peptides (AAB29544-
CC AAB29558) which modulate the adhesion of a target cell to a substrate.
CC The invention also encompasses substrates and devices treated with a
CC peptide of the invention; compositions comprising a peptide of the
CC invention for in vivo use; and analogues, fragments and chemical
CC derivatives of the peptides of the invention. The peptides are useful for
CC modulating the adhesion to a substrate of target cells such as
CC endothelial cells, fibroblasts, macrophages, neutrophils or
CC myofibroblasts. The substrate may be a substrate that is found in the
CC body of a patient, e.g., collagen or hyaluronic acid, or may be a
CC synthetic substrate e.g., a polyvinyl surface, titanium or PGA. The
CC peptides are useful for regulating vessel growth during wound healing
CC and/or in the treatment of damage resulting from vascular disease; for
CC inhibiting or preventing cellular apoptosis; in the treatment of
CC fibrosis, in particular in the clearing of debris; to minimise wound
CC contraction, thereby reducing keloid tissue formation and scarring; and
CC as anti clotting agents. The peptides also have an immunomodulatory
CC effect, and an antibacterial effect by adhering to neutrophils.
CC Additionally, peptides of the invention have an anticancer effect by
CC competing for alpha-v-beta3 integrin binding on the cell surface, and an
CC antitumorigenic effect by having anti-CD44 activity. The peptides are
CC useful for stimulating and/or enhancing cell attachment to polymer
CC scaffolds, to enhance tissue growth and for coating medical devices,
CC including prostheses and implants (e.g., vascular implants). The present
CC sequence represents a specifically claimed adhesion modulatory peptide of
CC the invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 22; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 3
ABP58757
ID ABP58757 standard; peptide; 4 AA.
XX
AC ABP58757;
XX
DT 04-APR-2003 (first entry)
XX
DE Macrophage recruitment inhibitor peptide, SEQ ID NO:7.
XX
KW Minicell display; random peptide library; drug discovery; screening;
KW fusion protein; outer membrane protein;
KW Rickettsia rickettsii 17K antigen; bioactive peptide;
KW macrophage recruitment inhibitor; osteopontin; C5a; fibronectin.
XX
OS Synthetic.
XX
PN WO200272759-A2.
XX
PD 19-SEP-2002.
XX
PF 06-MAR-2002; 2002WO-US06921.
XX
PR 07-MAR-2001; 2001US-274039P.
PR 20-JUL-2001; 2001US-306946P.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX
DR WPI; 2003-046738/04.
XX
PT Minicell display method for screening peptide libraries for peptide

PT candidates that can bind and modulate a particular biological process,
PT comprises identifying minicell hosts bound to the binding partner -
XX
PS Claim 22; Page 36; 66pp; English.
XX
CC The invention relates to minicell display methods for the generation
CC and screening of random peptide libraries for peptide candidates which
CC can bind to and modulate the activity of target molecules involved in
CC biological processes. The methods involve expressing a peptide library as
CC a fusion protein with a bacterial outer membrane protein (preferably the
CC 17K antigen of Rickettsia rickettsii) in an anucleate bacterially-derived
CC minicell host, and then contacting the minicell hosts with a target
CC molecule. Minicells comprising a library peptide which has bound to the
CC target molecule are separated from the unbound minicells, and the
CC peptides of interest analysed. The invention also encompasses methods for
CC making a minicell DNA library, increasing the diversity of a minicell DNA
CC library by in vivo mutagenesis, screening a minicell DNA library, and
CC purifying minicells. The methods of the invention are useful for
CC identifying minicell hosts bound to a target molecule and may be used in
CC drug discovery. The method increases the size of the peptides to be
CC screened and the diversification of the library to be screened, thereby
CC increasing the number of potential peptides that can modulate a
CC particular biological response or mechanism. Sequences ABP58757-ABP58803
CC represent bioactive peptides identified using the minicell display
CC method of the invention. The present sequence was characterised as
CC an inhibitor of macrophage recruitment by osteopontin, C5a and
CC fibronectin.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 24; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 4
AAB28661
ID AAB28661 standard; peptide; 6 AA.
XX
AC AAB28661;
XX
DT 13-FEB-2001 (first entry)
XX
DE Chemotactic peptide pepJ.
XX
KW Chemotactic; osteopontin; vulnerary; antiarthritic; antipsoriatic;
KW cytostatic; antitumour; antiinflammatory; osteopathic;
KW wound healing; cell migration; chemotaxis; atherosclerosis; cancer;
KW angiogenic-associated disease; arthritis; psoriasis; haemangioma;
KW ocular neovascularisation; cell apoptosis; nitrous oxide production;
KW inflammation; osteoporosis; immune disease.
XX
OS Mammalia.
OS Synthetic.
XX
PN WO200063247-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-US10344.
XX
PR 15-APR-1999; 99US-0129764.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX
DR WPI; 2000-687159/67.
XX
PT New osteopontin-derived chemotactic and inhibitory peptides, useful for
PT promoting scarless wound healing, modulating cellular chemotaxis,
PT treating formation of atherosclerotic plaques and preventing metastasis
PT -
XX
PS Claim 14; Page 42; 54pp; English.
XX
CC The present sequence is an osteopontin-derived chemotactic peptide.
CC Such chemotactic peptides are useful for promoting scarless wound
CC healing, modulating chemotaxis and promoting cell migration to a target
CC site in a cell of a subject. They are also used for modulating cellular
CC chemotaxis in a mammalian cell such as smooth muscle cell, a macrophage,
CC an endothelial cell, a vascular cell and a tumourigenic cell. They are
CC useful for treating the formation of atherosclerotic plaques in a
CC subject. The peptides are used for preventing metastasis, treating an
CC angiogenic-associated disease such as arthritis, psoriasis, haemangioma,
CC tumour metastasis or ocular neovascularisation. They are also used for
CC activating cell apoptosis, for modulating nitrous oxide production and
CC for inducing chemotaxis. The peptides are useful for diagnosing, treating
CC and preventing tumour metastasis, inflammation, osteoporosis and immune
CC diseases. They can also be used to enhance an immune response by
CC attracting macrophages.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 20; DB 21; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 2 VLEP 5

RESULT 5
ABP58014
ID ABP58014 standard; Peptide; 7 AA.
XX
AC ABP58014;
XX
DT 11-FEB-2003 (first entry)
XX
DE Prostate cancer marker protein peptide fragment.
XX
KW Prostate cancer; marker; vitamin D binding protein; VDBP; human;
KW diagnosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200275314-A2.
XX
PD 26-SEP-2002.
XX
PF 30-NOV-2001; 2001WO-US45031.
XX
PR 30-NOV-2000; 2000US-250284P.
PR 08-NOV-2001; 2001US-344948P.
XX
PA (MATR-) MATRITECH INC.
XX
PI Hlavaty J, Briggman JV;
XX
DR WPI; 2003-067369/06.
XX
PT Diagnosing or treating prostrate cancer by detecting in a sample
PT isolated from the individual the presence of prostate cancer-associated
PT protein -
XX
PS Claim 1; Page 41; 63pp; English.
XX
CC The present sequence is that of a peptide fragment of a novel human
CC 50.8 kDa prostate cancer-associated protein that has been identified
CC as a highly effective marker for prostate cancer. The novel protein

```
CC   includes a polypeptide that is related to human serum vitamin D
CC   binding protein (VDBP, see ABP58017).  The present peptide
CC   corresponds to amino acids 364-370 of this VDBP allele.  It is one of
CC   a series of peptides (see ABP58005-16) that distinguish VDBP-related
CC   proteins from other proteins, or which may be characterised as
CC   binding specifically to an anti-VDBP antibody.  VDBP-related proteins
CC   are detectable at a higher concentration in serum from a mammal, e.g.
CC   a human, with prostate cancer relative to serum from a healthy
CC   mammal and can therefore be used as prostate cancer markers.  They
CC   permit the rapid detection, preferably before metastases occur, of
CC   prostate cancer.  A target prostate cancer-associated protein may be
CC   detected using a labelled antibody capable of binding specifically
CC   to the protein.  Prostate cancer-associated proteins, and nucleic
CC   acids encoding them, are also useful as targets for treating
CC   prostate cancer, and as indicators for monitoring the efficiency of
CC   prostate cancer therapy.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 20;  DB 24;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          1 VLEP 4


RESULT 6
AAW06332
ID   AAW06332 standard; peptide; 8 AA.
XX
AC   AAW06332;
XX
DT   17-JAN-1997  (first entry)
XX
DE   N-terminus of receptor-type tyrosine kinase protein ligand.
XX
KW   Receptor-type tyrosine kinase; ligand; coomassie staining; PAS staining;
KW   N-terminus; human.
XX
OS   Homo sapiens.
XX
PN   JP08188596-A.
XX
PD   23-JUL-1996.
XX
PF   13-JAN-1995;   95JP-0003677.
XX
PR   09-NOV-1994;   94JP-0275411.
PR   19-OCT-1994;   94JP-0253848.
XX
PA   (ASAH ) ASAHI KASEI KOGYO KK.
XX
DR   WPI; 1996-388601/39.
XX
PT   New ligand for receptor type tyrosine kinase - has mol.wt. 22-25
PT   kilo:dalton(s) and is positive for Coomassie and PAS staining
XX
PS   Claim 3; Page 43; 51pp; Japanese.
XX
CC   This sequence represents the N-terminus of a receptor-type tyrosine
CC   kinase receptor binding ligand of the invention (see AAW06333 and
CC   AAW06334). The ligands of the invention recognise the fragment of the
CC   receptor type kinase receptor represented by AAW06330 (see AAW06331 for
CC   full length sequence).  The proteins of the invention have a molecular
CC   weight of 23500 (plus or minus 1500) Da, and are positive for Coomassie
CC   staining and PAS staining.  The protein is a new ligand of receptor-type
CC   tyrosine kinases, and can be prepared by standard recombinant
CC   techniques.
XX
SQ   Sequence   8 AA;

  Query Match             100.0%;  Score 20;  DB 17;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          4 VLEP 7


RESULT 7
AAR94654
ID   AAR94654 standard; Protein; 8 AA.
XX
AC   AAR94654;
XX
DT   18-OCT-1996  (first entry)
XX
DE   Fragment of receptor type tyrosine kinase (TK) ligand.
XX
KW   Receptor type tyrosine kinase; TK; ligand; differentiation;
KW   haematopoietic stem cell; tyrosine; bone marrow; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO9611212-A1.
XX
PD   18-APR-1996.
XX
PF   09-OCT-1995;   95WO-JP02069.
XX
PR   22-DEC-1994;   94JP-0320712.
PR   07-OCT-1994;   94JP-0244433.
PR   26-OCT-1994;   94JP-0262882.
XX
PA   (ASAH ) ASAHI KASEI KOGYO KK.
XX
PI   Ohno M,  Sakano S;
XX
DR   WPI; 1996-209809/21.
XX
PT   Ligand peptide binding to receptor-type tyrosine kinase - enhances
PT   intracellular tyrosine phosphorylation, useful for investigation of
PT   undifferentiated blood cell behaviour
XX
PS   Disclosure; Page 162; 193pp; Japanese.
XX
CC   A ligand polypeptide which binds to the extracellular part of a
CC   specific receptor-type tyrosine kinase and induces phosphorylation
CC   of tyrosine within the cell can be used in the study of the
CC   differentiation of blood cells such as the haematopoietic stem
CC   cells; of disease processes such as leukaemia, and of the biology of
CC   bone marrow transplantation.  The ligand plays a role in the
CC   differentiation process and the specific ligand target is expressed
CC   in undifferentiated blood cells.
XX
SQ   Sequence   8 AA;

  Query Match             100.0%;  Score 20;  DB 17;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          4 VLEP 7


RESULT 8
ABP99157
ID   ABP99157 standard; Peptide; 8 AA.
XX
AC   ABP99157;
XX
DT   18-MAR-2003  (first entry)
```

XX
DE ErbB2 cell overexpression EOPI-130 peptide SEQ ID NO:203.
XX
KW ErbB2; cancer; oncogene; ErbB2 overexpression-associated protein isoform;
KW EOPI; EOF; ErbB2 overexpression feature; cytostatic; vaccine;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200290991-A2.
XX
PD 14-NOV-2002.
XX
PF 02-MAY-2002; 2002WO-GB02047.
XX
PR 03-MAY-2001; 2001GB-0010886.
PR 23-NOV-2001; 2001GB-0028183.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Page MJ;
XX
DR WPI; 2003-103531/09.
XX
PT Diagnosing and treating ErbB2-related cancer, comprises generating
PT ErbB2 Overexpression Features (EOFs) from test samples from a subject
PT by electrophoresis, and comparing the EOFs in the sample with a
PT predetermined reference range -
XX
PS Claim 3; Page 23; 106pp; English.
XX
CC The present invention describes a method for screening or diagnosing
CC ErbB2-related cancer. The method comprises generating ErbB2
CC overexpression features (EOFs) from test samples of body fluid from the
CC subject by electrophoresis, and comparing the EOFs in the test sample
CC with that from normal subjects or with an expression reference feature
CC (ERF) in the test sample. Also described: (1) an antibody capable of
CC immunospecific binding to an ErbB2 overexpression protein isoform
CC (EOPI); (2) pharmaceutical compositions comprising an EOPI, a nucleic
CC acid encoding an EOPI, an amount of the above antibody or its fragment,
CC and a carrier; (3) a kit comprising one or more antibodies and/or EOPIs
CC cited above, other reagents and instructions for use; (4) methods of
CC treating or preventing ErbB2-related cancer; (5) methods of screening
CC for or identifying agents that interact with or modulate the expression
CC or activity of, one or more EOPIs, EOPI fragment, EOPI-related
CC polypeptides, or EOPI-fusion proteins; (6) a method for modulating the
CC activity of one or more of the ErbB2 EOPIs, comprising administering to
CC a subject an agent identified by the method of (5); and (7) a method for
CC identifying targets for therapeutic modulation of ErbB2-related cancer.
CC EOPIs have cytostatic activity and can be used in vaccines and gene
CC therapy. The method is useful in screening, diagnosing, preventing or
CC treating ErbB2-related cancer, determining the stage or severity of
CC ErbB2-related cancer, identifying a subject at risk of developing
CC ErbB2-related cancer, monitoring the effect of therapy administered to
CC a subject with ErbB2-related cancer, and for drug screening or drug
CC development. The kit is useful in carrying out the above methods.
CC ABP98940 to ABP99206 represent specifically claimed EOPIs from the
CC present invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 24; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 2 VLEP 5

RESULT 9
AAR84629
ID AAR84629 standard; Peptide; 9 AA.
XX
AC AAR84629;
XX
DT 02-JUN-1996 (first entry)
XX
DE Wheat acetyl-Coenzyme-A-carboxylase peptide sequence.
XX
KW Acetyl Coenzyme A carboxylase; ACCase; plasmid pK111;
KW transgenic plant; modified oil content;
KW polyhydroxyalkanoate polymer; herbicide resistance;
KW monocot crop improvement.
XX
OS Triticum aestivum.
XX
PN WO9529246-A1.
XX
PD 02-NOV-1995.
XX
PF 21-APR-1994; 94WO-GB00846.
XX
PR 21-APR-1994; 94WO-GB00846.
XX
PA (ZENE) ZENECA LTD.
XX
PI Bright SWJ, Elborough KM, Fentem PA, Slabas AR;
XX
DR WPI; 1995-382994/49.
XX
PT DNA encoding acetyl Coenzyme A carboxylase - used for developing
PT plants with controlled expression of ACCase, e.g. for controlling
PT fatty acid synthesis.
XX
PS Disclosure; Fig 3; 61pp; English.
XX
CC Wheat acetyl Coenzyme A carboxylase (ACCase) was partially
CC purified from wheat germ and a dominant 220 KDa band was identified
CC as ACCase by column chromatography and SDS-PAGE. After
CC purification in the gel, the protein was hydrolysed using
CC Endoproteinase LysC, and resulting peptides were purified by
CC electrophoresis and loaded onto an ABI 477A pulse liquid
CC protein sequencer. Sequence data of 4 peptides (this peptide and
CC peptides AAR84625, AAR84627 and AAR84631) were compared with plasmid
CC pK111-deduced amino acid sequences (AAR84630, AAR84626, AAR84628 and
CC AAR84632) so as to authenticate pK111 (see AAT04948) as wheat ACCase
CC partial cDNA. More specifically, this peptide corresponds to
CC amino acids 319-327 of the pK111 deduced AA sequence AAR84619
CC (i.e. AAR84630). Partial cDNA clone pK111 (NCIB 40553) can be
CC used to create a sense/antisense expression cassette to
CC transform rape and other oilseed plants (canola, soybean,
CC sunflower) to downregulate production of the ACCase enzyme.
CC The transgenic plants have a lower or a modified oil content.
CC Down-regulation of oil synthesis can be used to divert the
CC substrate, acetyl Coenzyme A, into synthesis of alternative
CC storage materials (starch, protein or novel polymers e.g.
CC polyhydroxyalkanoates). Full-length ACCase clones can be used
CC to create transgenic plants over-expressing ACCase, and
CC therefore with increased oil content. ACCase over-expression in
CC monocot plants such as wheat, barley, maize and rice, which are
CC normally sensitive to herbicides, results in
CC aryloxyphenoxy-propionate and alkylketone herbicide resistance in
CC the transgenic plants (dicots are normally resistant to these
CC herbicides).
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 20; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 2 VLEP 5

```
RESULT 10
AAY07169
ID   AAY07169 standard; Peptide; 10 AA.
XX
AC   AAY07169;
XX
DT   02-JUL-1999  (first entry)
XX
DE   HLA binding peptide.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US14679.
XX
PR   22-JUN-1998;   98US-0102322.
PR   17-JUL-1997;   97US-0896164.
PR   10-OCT-1997;   97US-0061599.
PR   10-OCT-1997;   97US-0061765.
PR   10-OCT-1997;   97US-0948705.
PR   11-OCT-1997;   97GB-0021697.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen Y,  Gout I,  Gure A,  O'Hare M,  Obata Y,  Old LJ;
PI   Pfreundschuh M,  Sahin U,  Scanlan MJ,  Stockert E;
PI   Tureci O;
XX
DR   WPI; 1999-132448/11.
XX
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
XX
PS   Example 16; Page 762; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence   10 AA;

  Query Match             100.0%;  Score 20;  DB 20;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          5 VLEP 8


RESULT 11
AAY07127
ID   AAY07127 standard; Peptide; 10 AA.
XX
AC   AAY07127;
XX
DT   02-JUL-1999  (first entry)
XX
DE   HLA binding peptide.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US14679.
XX
PR   22-JUN-1998;   98US-0102322.
PR   17-JUL-1997;   97US-0896164.
PR   10-OCT-1997;   97US-0061599.
PR   10-OCT-1997;   97US-0061765.
PR   10-OCT-1997;   97US-0948705.
PR   11-OCT-1997;   97GB-0021697.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen Y,  Gout I,  Gure A,  O'Hare M,  Obata Y,  Old LJ;
PI   Pfreundschuh M,  Sahin U,  Scanlan MJ,  Stockert E;
PI   Tureci O;
XX
DR   WPI; 1999-132448/11.
XX
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
XX
PS   Example 16; Page 755; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence   10 AA;

  Query Match             100.0%;  Score 20;  DB 20;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          5 VLEP 8


RESULT 12
ABP71117
ID   ABP71117 standard; peptide; 10 AA.
XX
AC   ABP71117;
XX
DT   14-APR-2003  (first entry)
```

```
XX
DE   Ced-3 protein CARD region fragment.
XX
KW   BTF3; cell death; apoptosis; basic transcription factor; cytostatic;
KW   nootropic; neuroprotective; antiparkinsonian; antiarteriosclerotic;
KW   antirheumatic; antiarthritic; gene therapy; CARD; ced-3.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "unknown"
XX
PN   WO200295001-A2.
XX
PD   28-NOV-2002.
XX
PF   21-MAY-2002; 2002WO-US16230.
XX
PR   21-MAY-2001; 2001US-292559P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Rothman JH,  Bloss T,  Witze E;
XX
DR   WPI; 2003-167228/16.
XX
PT   Inhibiting or increasing programmed cell death of a cell, for treating
PT   e.g. cancer, comprises upregulating or inhibiting, respectively, the
PT   expression or activity of basic transcription factor (BTF)3 or its
PT   homolog in the cell  -
XX
PS   Examples; Fig 2A; 84pp; English.
XX
CC   The invention relates to inhibiting or increasing programmed cell death
CC   of a cell. The method involves upregulating or inhibiting, respectively,
CC   the expression or activity of basic transcription factor (BTF)3 or its
CC   homologue in the cell. The BTF3 polypeptides and nucleic acids are useful
CC   for inhibiting or increasing programmed cell death. They are used for
CC   screening for an agent that increases or inhibits programmed cell death
CC   or pre-screening for an agent that modulates programmed cell death. The
CC   screened agent that increases or inhibits programmed cell death, is used
CC   for diagnosing or treating cancer or neurodegenerative diseases (e.g.
CC   amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease
CC   or multiple sclerosis), atherosclerosis, or rheumatoid arthritis.
CC   Sequences ABP71106-123 represent CARD regions of various CARD proteins.
XX
SQ   Sequence   10 AA;

  Query Match             100.0%;  Score 20;  DB 24;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          6 VLEP 9


RESULT 13
AAM00583
ID   AAM00583 standard; Peptide; 14 AA.
XX
AC   AAM00583;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human transcription factor fragment SEQ ID NO: 1131.
XX
KW   Human; single nucleotide polymorphism; SNP; paternity test;
KW   forensic test; aberrant protein expression.
XX
OS   Homo sapiens.
XX
PN   WO200151670-A2.
XX
PD   19-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-US00322.
XX
PR   07-JAN-2000; 2000US-0174962.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach MD;
XX
DR   WPI; 2001-451871/48.
DR   N-PSDB; AAH89700.
XX
PT   Isolated human polynucleotides containing single nucleotide
PT   polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
PT   infection and diabetes -
XX
PS   Disclosure; Page 427; 475pp; English.
XX
CC   The present invention relates to human nucleic acids containing single
CC   nucleotide polymorphisms (SNPs). These can be used in forensic and
CC   paternity tests, and to aid in the treatment of diseases associated with
CC   aberrant protein expression, including cancer, amyloidosis, diabetes,
CC   Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
CC   glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
CC   meningitis, muscular disorders, dementia, neurological diseases, tuberous
CC   sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
CC   osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
CC   autoimmunity. The present sequence is a pepttide encoded by a
CC   polymorphism-containing oligonucleotide fragment of the invention.
XX
SQ   Sequence   14 AA;

  Query Match             100.0%;  Score 20;  DB 22;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          5 VLEP 8


RESULT 14
AAB08085
ID   AAB08085 standard; peptide; 17 AA.
XX
AC   AAB08085;
XX
DT   04-DEC-2000  (first entry)
XX
DE   Amino acid sequence of a helper T cell epitope from CDV.
XX
KW   T helper cell epitope; CDV; immune response; canine vaccine.
XX
OS   Canine distemper virus.
XX
PN   WO200046390-A1.
XX
PD   10-AUG-2000.
XX
PF   07-FEB-2000; 2000WO-AU00070.
XX
PR   05-FEB-1999;   99AU-0008533.
PR   04-AUG-1999;   99AU-0002013.
XX
PA   (UYME ) UNIV MELBOURNE.
PA   (CSLC-) CSL LTD.
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
```

PI Jackson DC, Souravi G, Walker J;
XX
DR WPI; 2000-532904/48.
XX
PT Novel T helper cell epitopes derived from canine distemper virus useful
PT for preparation of canine vaccines -
XX
PS Claim 1; Page 28; 54pp; English.
XX
CC AAB08076-B08101 represent T helper cell epitopes, derived from canine
CC distemper virus (CDV). Compositions comprising these T cell helper
CC epitopes are useful for inducing an immune response in an animal. The
CC epitopes are useful as components of animal, in particular, canine
CC vaccines, either simply as synthetic peptide based vaccines and as
CC additions to vaccines containing more complex antigens.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 20; DB 21; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 13 VLEP 16

RESULT 15
AAB08086
ID AAB08086 standard; peptide; 17 AA.
XX
AC AAB08086;
XX
DT 04-DEC-2000 (first entry)
XX
DE Amino acid sequence of a helper T cell epitope from CDV.
XX
KW T helper cell epitope; CDV; immune response; canine vaccine.
XX
OS Canine distemper virus.
XX
PN WO200046390-A1.
XX
PD 10-AUG-2000.
XX
PF 07-FEB-2000; 2000WO-AU00070.
XX
PR 05-FEB-1999; 99AU-0008533.
PR 04-AUG-1999; 99AU-0002013.
XX
PA (UYME) UNIV MELBOURNE.
PA (CSLC-) CSL LTD.
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Jackson BC, Souravi G, Walker J;
XX
DR WPI; 2000-532904/48.
XX
PT Novel T helper cell epitopes derived from canine distemper virus useful
PT for preparation of canine vaccines -
XX
PS Claim 1; Page 28; 54pp; English.
XX
CC AAB08076-B08101 represent T helper cell epitopes, derived from canine
CC distemper virus (CDV). Compositions comprising these T cell helper
CC epitopes are useful for inducing an immune response in an animal. The
CC epitopes are useful as components of animal, in particular, canine
CC vaccines, either simply as synthetic peptide based vaccines and as
CC additions to vaccines containing more complex antigens.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 20; DB 21; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 6 VLEP 9

Search completed: November 25, 2003, 13:59:06
Job time : 44 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 13:58:16 ; Search time 22 Seconds
(without alignments)
7.693 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	9	4	US-08-737-109-5	Sequence 5, Appli
2	20	100.0	19	3	US-08-928-213B-35	Sequence 35, Appl
3	20	100.0	19	3	US-09-001-984C-72	Sequence 72, Appl
4	20	100.0	19	4	US-09-396-347F-72	Sequence 72, Appl
5	20	100.0	25	1	US-08-127-909-20	Sequence 20, Appl
6	20	100.0	25	2	US-08-457-353-20	Sequence 20, Appl
7	20	100.0	25	3	US-09-011-525-5	Sequence 5, Appli
8	20	100.0	26	1	US-07-942-245-281	Sequence 281, App
9	20	100.0	26	4	US-08-525-539A-21	Sequence 21, Appl
10	20	100.0	34	2	US-08-967-508-1	Sequence 1, Appli
11	20	100.0	34	3	US-08-967-506-1	Sequence 1, Appli
12	20	100.0	34	5	PCT-US94-02552-1	Sequence 1, Appli
13	20	100.0	35	3	US-09-082-279B-504	Sequence 504, App
14	20	100.0	35	4	US-09-315-304B-504	Sequence 504, App
15	20	100.0	35	4	US-09-834-784-504	Sequence 504, App
16	20	100.0	40	4	US-09-314-268-106	Sequence 106, App
17	20	100.0	65	6	5320958-17	Patent No. 5320958
18	20	100.0	70	4	US-09-134-001C-4579	Sequence 4579, Ap
19	20	100.0	74	4	US-09-107-532A-4364	Sequence 4364, Ap
20	20	100.0	75	3	US-08-928-383B-13	Sequence 13, Appl
21	20	100.0	82	4	US-09-655-270A-25	Sequence 25, Appl
22	20	100.0	84	4	US-09-464-535-18	Sequence 18, Appl
23	20	100.0	90	4	US-09-252-991A-29776	Sequence 29776, A
24	20	100.0	93	3	US-09-208-804-3	Sequence 3, Appli
25	20	100.0	93	3	US-08-801-743-3	Sequence 3, Appli
26	20	100.0	98	3	US-08-478-097A-5	Sequence 5, Appli
27	20	100.0	98	4	US-09-496-398-5	Sequence 5, Appli
28	20	100.0	100	2	US-08-963-601-2	Sequence 2, Appli
29	20	100.0	100	4	US-08-735-848-2	Sequence 2, Appli
30	20	100.0	101	4	US-09-216-393B-114	Sequence 114, App
31	20	100.0	101	4	US-09-216-393B-268	Sequence 268, App
32	20	100.0	105	3	US-08-488-551B-639	Sequence 639, App
33	20	100.0	105	4	US-09-732-210-1021	Sequence 1021, Ap
34	20	100.0	106	4	US-09-198-452A-6845	Sequence 6845, Ap
35	20	100.0	108	3	US-08-388-353-639	Sequence 639, App
36	20	100.0	109	4	US-09-328-352-4485	Sequence 4485, Ap
37	20	100.0	111	4	US-09-198-452A-1159	Sequence 1159, Ap
38	20	100.0	112	3	US-08-927-433-7	Sequence 7, Appli
39	20	100.0	114	1	US-08-481-377-27	Sequence 27, Appl
40	20	100.0	114	2	US-08-491-835-25	Sequence 25, Appl
41	20	100.0	114	3	US-09-153-733A-27	Sequence 27, Appl
42	20	100.0	114	3	US-08-946-092A-25	Sequence 25, Appl
43	20	100.0	114	3	US-09-172-062-25	Sequence 25, Appl
44	20	100.0	114	4	US-09-301-520D-25	Sequence 25, Appl
45	20	100.0	114	4	US-09-389-705-27	Sequence 27, Appl

ALIGNMENTS

```
RESULT 1
US-08-737-109-5
; Sequence 5, Application US/08737109
; Patent No. 6455688
;  GENERAL INFORMATION:
;    APPLICANT:  SLABAS, Antoni Ryszard
;    APPLICANT:  ELBOROUGH, Kieran Michael
;    APPLICANT:  BRIGHT, Simon William Jonathan
;    APPLICANT:  FENTEM, Philip Anthony
;    TITLE OF INVENTION:  Plant Gene Specifying Acetyl Coenzyme A
;    TITLE OF INVENTION:  Carboxylase and Transformed Plants Containing Same
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pillsbury Madison & Sutro, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/737,109
;      FILING DATE:  21-OCT-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB94/00846
;      FILING DATE:  02-MAY-1994
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Avena sativa
US-08-737-109-5

  Query Match           100.0%;  Score 20;  DB 4;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          2 VLEP 5
```

```
RESULT 2
US-08-928-213B-35
; Sequence 35, Application US/08928213B
; Patent No. 6238905
;  GENERAL INFORMATION:
;       APPLICANT: McHenry, Charles S.
;                  Seville, Mark
;                  Cull, Millard G.
;       TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III
;                           HOLOENZYME
;       NUMBER OF SEQUENCES: 195
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: MEDLEN & CARROLL, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/928,213B
;            FILING DATE: 12-Sep-1997
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: MacKnight, Kamrin T.
;            REGISTRATION NUMBER: 38,230
;            REFERENCE/DOCKET NUMBER: ENZYCO-02550
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-705-8410
;            TELEFAX: 415-397-8338
;  INFORMATION FOR SEQ ID NO: 35:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 19 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: not relevant
;            TOPOLOGY: not relevant
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-08-928-213B-35

  Query Match             100.0%;  Score 20;  DB 3;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          1 VLEP 4


RESULT 3
US-09-001-984C-72
; Sequence 72, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
;  APPLICANT:  Laal, Suman
;  APPLICANT:  Zolla-Pazner, Susan
;  APPLICANT:  Belisle, John T
;  TITLE OF INVENTION:  EARLY DETECTION OF MYCOBACTERIAL DISEASE
;  FILE REFERENCE:  NYU-011
;  CURRENT APPLICATION NUMBER: US/09/001,984C
;  CURRENT FILING DATE:  1997-12-31
;  PRIOR APPLICATION NUMBER:  60/034,003
;  PRIOR FILING DATE:  1996-12-31
;  NUMBER OF SEQ ID NOS:  106
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 72
;  LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-001-984C-72

  Query Match             100.0%;  Score 20;  DB 3;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         15 VLEP 18


RESULT 4
US-09-396-347F-72
; Sequence 72, Application US/09396347F
; Patent No. 6506384
; GENERAL INFORMATION:
;  APPLICANT: Laal, Suman
;  APPLICANT:  Zolla-Pazner, Susan
;  APPLICANT:  Belisle, John T
;  TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
;  FILE REFERENCE: 32004-169276
;  CURRENT APPLICATION NUMBER: US/09/396,347F
;  CURRENT FILING DATE:  1999-09-14
;  PRIOR APPLICATION NUMBER: 09/001,984
;  PRIOR FILING DATE: 1997-12-31
;  NUMBER OF SEQ ID NOS: 106
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
;  LENGTH: 19
;  TYPE: PRT
;  ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-396-347F-72

  Query Match             100.0%;  Score 20;  DB 4;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         15 VLEP 18


RESULT 5
US-08-127-909-20
; Sequence 20, Application US/08127909
; Patent No. 5436228
;  GENERAL INFORMATION:
;    APPLICANT:  Postlethwaite, Arnold E.
;    APPLICANT:  Seyer, Jerome
;    APPLICANT:  Kang, Andrew
;    TITLE OF INVENTION:  CHEMOTACTIC WOUND HEALING PEPTIDES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/127,909
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9020Z
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-127-909-20

  Query Match           100.0%;  Score 20;  DB 1;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          4 VLEP 7


RESULT 6
US-08-457-353-20
; Sequence 20, Application US/08457353
; Patent No. 5824647
;  GENERAL INFORMATION:
;    APPLICANT:  Postlethwaite, Arnold E.
;    APPLICANT:  Seyer, Jerome
;    APPLICANT:  Kang, Andrew
;    TITLE OF INVENTION:  CHEMOTACTIC WOUND HEALING PEPTIDES
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/457,353
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9020Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-457-353-20

  Query Match           100.0%;  Score 20;  DB 2;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          4 VLEP 7


RESULT 7
US-09-011-525-5
; Sequence 5, Application US/09011525
; Patent No. 6020172
; GENERAL INFORMATION:
;  APPLICANT: BOTH, GERALD W.
;  TITLE OF INVENTION: GENE THERAPY USING OVINE ADENOVIRAL VECTORS
;  FILE REFERENCE: Gene Therapy Using Ovine Adenoviral Ve
;  CURRENT APPLICATION NUMBER: US/09/011,525
;  CURRENT FILING DATE: 1998-04-20
;  EARLIER APPLICATION NUMBER: PCT/AU96/00518
;  EARLIER FILING DATE: 1996-08-14
;  EARLIER APPLICATION NUMBER: AU PN4776
;  EARLIER FILING DATE: 1995-08-14
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Ovine adenovirus
US-09-011-525-5

  Query Match           100.0%;  Score 20;  DB 3;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          7 VLEP 10


RESULT 8
US-07-942-245-281
; Sequence 281, Application US/07942245
; Patent No. 5639641
;  GENERAL INFORMATION:
;    APPLICANT:  PEDERSEN, Jan T.
;    APPLICANT:  SEARLE,   Stephen M.J.
;    APPLICANT:  REES,     Anthony R.
;    APPLICANT:  ROGUSKA,  Michael A.
;    APPLICANT:  GUILD,    Braydon C.
;    TITLE OF INVENTION:  SURFACE RESIDUE VENEERING OF RODENT
;    TITLE OF INVENTION:  ANTIBODIES
;    NUMBER OF SEQUENCES:  522
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pensylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  HP 9000/700 Workstation
;      OPERATING SYSTEM:  UNIX
;      SOFTWARE:  In house
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/942,245
;      FILING DATE:  09-SEP-1992
;      CLASSIFICATION:  530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  281:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-942-245-281

  Query Match           100.0%;  Score 20;  DB 1;  Length 26;
  Best Local Similarity  100.0%;  Pred. No. 1.3e+02;
```

```
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          5 VLEP 8


RESULT 9
US-08-525-539A-21
; Sequence 21, Application US/08525539A
; Patent No. 6309636
;  GENERAL INFORMATION:
;    APPLICANT:  DO COUTO, FERNANDO J.R.
;    APPLICANT:  CERIANI, ROBERTO L.
;    APPLICANT:  PETERSON, JERRY A.
;    TITLE OF INVENTION:  RECOMBINANT PEPTIDES DERIVED FROM THE
;    TITLE OF INVENTION:  Mc3 ANTI-BA46 ANTIBODY, METHODS OF USE THEREOF, AND
;    TITLE OF INVENTION:  METHODS OF HUMANIZING ANTIBODY PEPTIDES
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,539A
;      FILING DATE:  14-SEP-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DYLAN, TYLER
;      REGISTRATION NUMBER:  37,612
;      REFERENCE/DOCKET NUMBER:  27633-20001.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-525-539A-21

  Query Match             100.0%;  Score 20;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          1 VLEP 4


RESULT 10
US-08-967-508-1
; Sequence 1, Application US/08967508
; Patent No. 5910570
;  GENERAL INFORMATION:
;    APPLICANT:  The Upjohn Company
;    APPLICANT:  FOR U.S. PURPOSES ONLY:  Elhammer, Ake P. and Homa, Fred L.
;    TITLE OF INVENTION:  A Cloned DNA Encoding a UDP-GaLNAc:
;    TITLE OF INVENTION:  Polypeptide, N-Acetylgalactosaminyltransferase
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pharmacia & Upjohn Company, Intellectual
;      ADDRESSEE:  Property Legal Services
;      STREET:  301 Henrietta Street
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,508
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/602,830
;      FILING DATE:  13 No. 5910570ember 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Darnley Jr., James D.
;      REGISTRATION NUMBER:  33,673
;      REFERENCE/DOCKET NUMBER:  4755.P CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  616-833-2210
;      TELEFAX:  616-833-8897
;      TELEX:  224401
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-967-508-1

  Query Match             100.0%;  Score 20;  DB 2;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          7 VLEP 10


RESULT 11
US-08-967-506-1
; Sequence 1, Application US/08967506
; Patent No. 6096512
;  GENERAL INFORMATION:
;    APPLICANT:  The Upjohn Company
;    APPLICANT:  FOR U.S. PURPOSES ONLY:  Elhammer, Ake P. and Homa, Fred L.
;    TITLE OF INVENTION:  A Cloned DNA Encoding a UDP-GaLNAc:
;    TITLE OF INVENTION:  Polypeptide, N-Acetylgalactosaminyltransferase
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pharmacia & Upjohn Company, Intellectual
;      ADDRESSEE:  Property Legal Services
;      STREET:  301 Henrietta Street
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,506
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/602,830
;      FILING DATE:  13 No. 6096512ember 1995
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Darnley Jr., James D.
;      REGISTRATION NUMBER:  33,673
;      REFERENCE/DOCKET NUMBER:  4755.P CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  616-833-2210
;      TELEFAX:  616-833-8897
;      TELEX:  224401
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-967-506-1

  Query Match           100.0%;  Score 20;  DB 3;  Length 34;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          7 VLEP 10


RESULT 12
PCT-US94-02552-1
; Sequence 1, Application PC/TUS9402552
;  GENERAL INFORMATION:
;    APPLICANT:  Elhammer, Ake P.
;    APPLICANT:  Homa, Fred L.
;    TITLE OF INVENTION:  A Cloned DNA Encoding a UDP-GaLNAc:
;    TITLE OF INVENTION:  Polypeptide, N-Acetylgalactosaminyltransferase
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Upjohn Company, Corp. Intellectual
;      ADDRESSEE:  Property Law
;      STREET:  301 Henrietta Street
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/02552
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Darnley Jr., James D.
;      REGISTRATION NUMBER:  33,673
;      REFERENCE/DOCKET NUMBER:  4755.P CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  616-385-5210
;      TELEFAX:  616-385-6897
;      TELEX:  224401
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US94-02552-1

  Query Match           100.0%;  Score 20;  DB 5;  Length 34;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          7 VLEP 10


RESULT 13
US-09-082-279B-504
; Sequence 504, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
;  APPLICANT:  Barney, Shawn
;  APPLICANT:  Guthrie, Kelly
;  APPLICANT:  Merutka, Gene
;  APPLICANT:  Anwer, Mohmed
;  APPLICANT:  Lambert, Dennis
;  TITLE OF INVENTION:  HYBRID POLYPEPTIDES WITH ENHANCED
;  TITLE OF INVENTION:  PHARMACOKINETIC PROPERTIES
;  FILE REFERENCE:  7872-043
;  CURRENT APPLICATION NUMBER: US/09/082,279B
;  CURRENT FILING DATE:  1998-05-20
;  NUMBER OF SEQ ID NOS:  1515
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 504
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Core polypeptide
US-09-082-279B-504

  Query Match           100.0%;  Score 20;  DB 3;  Length 35;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         13 VLEP 16


RESULT 14
US-09-315-304B-504
; Sequence 504, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
;  APPLICANT: Barney, S.
;  APPLICANT:  Guthrie, K.
;  APPLICANT:  Merutka, G.
;  APPLICANT:  Anwer, M.
;  APPLICANT:  Lambert, D.
;  TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
;  TITLE OF INVENTION:  PROPERTIES
;  FILE REFERENCE: 7872-052
;  CURRENT APPLICATION NUMBER: US/09/315,304B
;  CURRENT FILING DATE:  1999-05-20
;  PRIOR APPLICATION NUMBER: 09/082,279
;  PRIOR FILING DATE: 1998-05-20
;  NUMBER OF SEQ ID NOS: 1667
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 504
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Core polypeptide
US-09-315-304B-504

  Query Match           100.0%;  Score 20;  DB 4;  Length 35;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         13 VLEP 16
```

```
RESULT 15
US-09-834-784-504
; Sequence 504, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
;  APPLICANT: Barney, Shawn
;  APPLICANT:  Guthrie, Kelly
;  APPLICANT:  Merutka, Gene
;  APPLICANT:  Anwer, Mohmed
;  APPLICANT:  Lambert, Dennis
;  TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
;  TITLE OF INVENTION:  PHARMACOKINETIC PROPERTIES
;  FILE REFERENCE: 7872-043
;  CURRENT APPLICATION NUMBER: US/09/834,784
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: 09/082,279
;  PRIOR FILING DATE: 1998-05-20
;  NUMBER OF SEQ ID NOS: 1515
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 504
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Core polypeptide
US-09-834-784-504

  Query Match             100.0%;  Score 20;  DB 4;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db         13 VLEP 16


Search completed: November 25, 2003, 14:01:26
Job time : 23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:05:07 ; Search time 21 Seconds
(without alignments)
18.318 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 86

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	12	60.0	3	3	A33802	thyrotropin-releas
2	11	55.0	4	2	I51049	metallothionein-A
3	10	50.0	4	2	E44823	synaptosomal-assoc
4	8	40.0	4	2	A02147	phagocytosis-stimu
5	7	35.0	3	3	RHTDTO	thyroliberin - Bom
6	7	35.0	3	3	RHPGT	thyroliberin - pig
7	7	35.0	3	3	RHSHT	thyroliberin - she
8	7	35.0	3	3	A92971	thyroliberin - eas
9	7	35.0	3	3	A43391	TRH-like tripeptid
10	7	35.0	3	3	I78890	tyrosine protein k
11	7	35.0	4	2	A32039	tyrosine-melanocyt
12	7	35.0	4	2	PL0140	carbon-monoxide de
13	7	35.0	4	2	A61300	22K superhelical D
14	7	35.0	4	2	I57745	D-mannonate hydrol
15	7	35.0	4	2	S53508	starvation-induced
16	7	35.0	4	2	S17255	ribosomal protein
17	7	35.0	4	2	A34626	RPCH-related neuro
18	7	35.0	4	2	PT0240	Ig heavy chain CRD
19	7	35.0	4	2	I54357	schwannomin - mous
20	7	35.0	4	2	PT0675	T-cell receptor be
21	6	30.0	4	2	A48360	gamma subunit of P
22	6	30.0	4	2	A27897	glucan 1,4-alpha-g
23	5	25.0	4	2	A41890	protein D - Escher
24	5	25.0	4	2	B43848	cell surface adhes
25	5	25.0	4	2	I40505	hypothetical prote
26	5	25.0	4	2	PT0677	T-cell receptor be
27	5	25.0	4	2	A26209	protein-glutamine
28	5	25.0	4	2	S55238	pallidipin - assas
29	4	20.0	3	3	PQ0010	angiotensin-conver
30	4	20.0	3	3	T13892	cytochrome-c oxida
31	4	20.0	4	2	S18401	thyroglobulin - do
32	4	20.0	4	2	A37832	phenol 2-monooxyge
33	4	20.0	4	2	T46627	hypothetical prote
34	4	20.0	4	2	T30569	hypothetical prote
35	4	20.0	4	2	I38888	COI intron 16 prot
36	4	20.0	4	2	A35779	neuropeptide Antho
37	4	20.0	4	2	A53284	T-cell receptor be
38	4	20.0	4	2	PT0645	T-cell receptor be
39	4	20.0	4	2	PT0721	T-cell receptor be
40	4	20.0	4	2	A40135	branched-chain-ami
41	4	20.0	4	2	S47552	ubiquitin - rat
42	4	20.0	4	2	S09478	globulin IV alpha
43	3	15.0	3	3	B23751	spinal cord peptid
44	3	15.0	3	3	S13894	histidinol dehydro
45	3	15.0	4	2	I40870	phospholipase C (E

ALIGNMENTS

RESULT 1
A33802
thyrotropin-releasing hormone-like peptide - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A33802
R;Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.
J. Biol. Chem. 264, 7788-7791, 1989
A;Title: A novel peptide, pyroglutamylglutamylproline amide, in the rabbit prostate com
A;Reference number: A33802; MUID:89255196; PMID:2498305
A;Accession: A33802
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3 <COC>
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 60.0%; Score 12; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EP 4
||
Db 2 EP 3

RESULT 2
I51049
metallothionein-A - rainbow trout (fragment)
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51049
R;Olsson, P.E.; Kling, P.; Erkell, L.J.; Kille, P.
Eur. J. Biochem. 230, 344-349, 1995
A;Title: Structural and functional analysis of the rainbow trout (Oncorhyncus mykiss) m
A;Reference number: I51049; MUID:95324545; PMID:7601121
A;Accession: I51049
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <OLS>
A;Cross-references: EMBL:X80181; NID:g1019799; PIDN:CAA56466.1; PID:g4379328

Query Match 55.0%; Score 11; DB 2; Length 4;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
::|
Db 1 MDP 3

```
RESULT 3
E44823
synaptosomal-associated protein SNAP-25 peptide 1 - rabbit (fragment)
N;Alternate names: superprotein peptide 1
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 15-Jun-1996
C;Accession: E44823
R;Loewy, A.; Liu, W.S.; Baitinger, C.; Willard, M.B.
J. Neurosci. 11, 3412-3421, 1991
A;Title: The major 35S-methionine-labeled rapidly transported protein (superprotein) is
A;Reference number: A44823; MUID:92044785; PMID:1941090
A;Accession: E44823
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <LOE>
A;Experimental source: visual tissue
A;Note: sequence extracted from NCBI backbone (NCBIP:64247)
C;Keywords: membrane trafficking

  Query Match             50.0%;  Score 10;  DB 2;  Length 4;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLE 3
              ::|
Db          1 IME 3


RESULT 4
A02147
phagocytosis-stimulating peptide (tuftsin) - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 03-Feb-1994
C;Accession: A02147
R;Nishioka, K.; Constantopoulos, A.; Satoh, P.S.; Najjar, V.A.
Biochem. Biophys. Res. Commun. 47, 172-179, 1972
A;Title: The characteristics, isolation and synthesis of the phagocytosis stimulating pe
A;Reference number: A02147; MUID:72187087; PMID:4112769
A;Accession: A02147
A;Molecule type: protein
A;Residues: 1-4 <NIS>
A;Note: a peptide having the same structure, physical properties, and biological activit
R;Fidalgo, B.V.; Najjar, V.A.
Biochemistry 6, 3386-3392, 1967
A;Reference number: A37502; MUID:68091045; PMID:4169272
A;Contents: annotation; immunoglobulin class
C;Comment: An IgG (called leucokinin) binds reversibly to the cell membrane of neutrophi
n is essential for maximum stimulation of the phagocytic activity of neutrophils.
C;Superfamily: immunoglobulin C region; immunoglobulin homology

  Query Match             40.0%;  Score 8;  DB 2;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 EP 4
              :|
Db          2 KP 3


RESULT 5
RHTDTO
thyroliberin - Bombina orientalis
C;Species: Bombina orientalis
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A90919; A01415
R;Yasuhara, T.; Nakajima, T.
Chem. Pharm. Bull. 23, 3301-3303, 1975
A;Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.
A;Reference number: A90919; MUID:76138399; PMID:815011
A;Accession: A90919
A;Molecule type: protein
A;Residues: 1-3 <YAS>
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

  Query Match             35.0%;  Score 7;  DB 3;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 P 4
              |
Db          3 P 3


RESULT 6
RHPGT
thyroliberin - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A01415
R;Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.
Biochemistry 9, 1103-1106, 1970
A;Title: Structure of porcine thyrotropin releasing hormone.
A;Reference number: A90560; MUID:70136150; PMID:4984938
A;Accession: A01415
A;Molecule type: protein
A;Residues: 1-3 <NAI>
R;Boler, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.
Biochem. Biophys. Res. Commun. 37, 705-710, 1969
A;Title: The identity of chemical and hormonal properties of the thyrotropin releasing h
A;Reference number: A90167; MUID:70039904; PMID:4982117
A;Contents: annotation
A;Note: biological activities and Rf values (in 17 chromatographic systems) of the synth
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

  Query Match             35.0%;  Score 7;  DB 3;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 P 4
              |
Db          3 P 3


RESULT 7
RHSHT
thyroliberin - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A93750; A01415
R;Desiderio Jr., D.M.; Burgus, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.
Org. Mass Spectrom. 5, 221-228, 1971
A;Title: The elucidation of the primary structure of the hypothalamic thyroid stimulati
A;Reference number: A93750
A;Accession: A93750
A;Molecule type: protein
A;Residues: 1-3 <DES>
R;Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.
Nature 226, 321-325, 1970
A;Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.
A;Reference number: A93161; MUID:70163386; PMID:4985794
A;Contents: annotation
A;Note: physicochemical characteristics and biological activities of the natural and sy
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

  Query Match             35.0%;  Score 7;  DB 3;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
```

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 P 4
|
Db 3 P 3

RESULT 8
A92971
thyroliberin - eastern newt (tentative sequence)
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A92971; A01415
R;Grimm-Jorgensen, Y.; McKelvy, J.F.
J. Neurochem. 23, 471-478, 1974
A;Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) bra
A;Reference number: A92971; MUID:75035605; PMID:4214528
A;Accession: A92971
A;Molecule type: protein
A;Residues: 1-3 <GRI>
A;Note: a peptide with the chromatographic and electrophoretic characteristics of thyrol
stidine, or glutamic acid
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 35.0%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 P 4
|
Db 3 P 3

RESULT 9
A43391
TRH-like tripeptide - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A43391
R;Lackey, D.B.
J. Biol. Chem. 267, 17508-17511, 1992
A;Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroGlu-
A;Reference number: A43391; MUID:92388092; PMID:1517203
A;Accession: A43391
A;Molecule type: protein
A;Residues: 1-3 <LAC>
C;Keywords: amidated carboxyl end; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 35.0%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 P 4
|
Db 3 P 3

RESULT 10
I78890
tyrosine protein kinase - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: I78890
R;Chow, L.M.; Davidson, D.; Fournel, M.; Gosselin, P.; Lemieux, S.; Lyu, M.S.; Kozak, C.
Oncogene 9, 3437-3448, 1994
A;Title: Two distinct protein isoforms are encoded by ntk, a csk-related tyrosine protei
A;Reference number: I58407; MUID:95060800; PMID:7970703
A;Accession: I78890
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3 <RES>
A;Cross-references: GB:L33339; NID:g609536; PIDN:AAA64432.1; PID:g609538
C;Genetics:
A;Gene: p52ntk

Query Match 35.0%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 P 4
|
Db 2 P 2

RESULT 11
A32039
tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 18-Aug-2000
C;Accession: A32039
R;Horvath, A.; Kastin, A.J.
J. Biol. Chem. 264, 2175-2179, 1989
A;Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor
A;Reference number: A32039; MUID:89123285; PMID:2563371
A;Accession: A32039
A;Molecule type: protein
A;Residues: 1-4 <HOR>
A;Experimental source: brain
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end
F;4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.0%; Score 7; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 P 4
|
Db 2 P 2

RESULT 12
PL0140
carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydohydrogena
C;Species: Pseudomonas carboxydohydrogena
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C;Accession: PL0140
R;Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A;Title: Homology and distribution of CO dehydrogenase structural genes in carboxydotro
A;Reference number: PL0138; MUID:90055678; PMID:2818128
A;Accession: PL0140
A;Molecule type: protein
A;Residues: 1-4 <KRA>
C;Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large, m
C;Keywords: oxidoreductase

Query Match 35.0%; Score 7; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 P 4
|
Db 4 P 4

RESULT 13
A61300
22K superhelical DNA-binding protein - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999

```
C;Accession: A61300
R;Kishi, F.; Ebina, Y.; Miki, T.; Nakazawa, T.; Nakazawa, A.
J. Biochem. 92, 1059-1068, 1982
A;Title: Purification and characterization of a protein from Escherichia coli which form
A;Reference number: A61300; MUID:83082696; PMID:6294066
A;Accession: A61300
A;Molecule type: protein
A;Residues: 1-4 <KIS>
C;Comment: This protein resembles some of the histone-like protein of bacteria in amino
C;Keywords: DNA binding; monomer

  Query Match             35.0%;  Score 7;  DB 2;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LE 3
              :|
Db          1 ME 2


RESULT 14
I57745
D-mannonate hydrolase (uxuA) - Escherichia coli
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 08-Oct-1999
C;Accession: I57745
R;Blanco, C.; Ritzenthaler, P.; Kolb, A.
Mol. Gen. Genet. 202, 112-119, 1986
A;Title: The regulatory region of the uxuAB operon in Escherichia coli K12.
A;Reference number: I57745; MUID:86174344; PMID:3083215
A;Accession: I57745
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: EMBL:X03411; NID:g43300; PIDN:CAA27147.1; PID:g581254

  Query Match             35.0%;  Score 7;  DB 2;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LE 3
              :|
Db          1 ME 2


RESULT 15
S53508
starvation-induced ribonuclease - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C;Accession: S53508
R;Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.
Plant Mol. Biol. 27, 477-485, 1995
A;Title: cDNA structure and regulatory properties of a family of starvation-induced ribo
A;Reference number: S53506; MUID:95201242; PMID:7894013
A;Accession: S53508
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <KOE>

  Query Match             35.0%;  Score 7;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 P 4
              |
Db          2 P 2


Search completed: November 25, 2003, 14:08:22
Job time : 21 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:01:32 ; Search time 11 Seconds
(without alignments)
17.101 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8	40.0	4	1	TUFT_HUMAN	P01858 homo sapien
2	7	35.0	3	1	THYL_PIG	P01151 sus scrofa
3	7	35.0	4	1	DCML_PSECH	P19916 pseudomonas
4	7	35.0	4	1	RM01_YEAST	P36515 saccharomyc
5	5	25.0	4	1	EOSI_HUMAN	P02731 homo sapien
6	4	20.0	4	1	FAR3_HIRME	P42562 hirudo medi
7	4	20.0	4	1	FLRF_HIRME	P42561 hirudo medi
8	4	20.0	4	1	FLRN_ANTEL	P58707 anthopleura
9	3	15.0	3	1	LUXE_VIBFI	P24272 vibrio fisc
10	3	15.0	4	1	FYRI_ANTEL	P58706 anthopleura
11	2	10.0	4	1	ACH1_ACHFU	P35904 achatina fu
12	2	10.0	4	1	DCMS_PSECH	P19918 pseudomonas
13	2	10.0	4	1	FAR4_HIRME	P42563 hirudo medi
14	2	10.0	4	1	FMRF_MACNI	P01162 macrocallis
15	2	10.0	4	1	OCP1_OCTMI	P58648 octopus min
16	2	10.0	4	1	OCP3_OCTMI	P58649 octopus min
17	1	5.0	3	1	GRWM_HUMAN	P01157 homo sapien
18	1	5.0	4	1	FFKA_ANTEL	P58705 anthopleura

ALIGNMENTS

RESULT 1
TUFT_HUMAN
ID TUFT_HUMAN STANDARD; PRT; 4 AA.
AC P01858;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phagocytosis-stimulating peptide (Tuftsin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=72187087; PubMed=4112769;
RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
RT "The characteristics, isolation and synthesis of the phagocytosis
RT stimulating peptide tuftsin.";
RL Biochem. Biophys. Res. Commun. 47:172-179(1972).
RN [2]
RP IMMUNOGLOBULIN CLASS.
RX MEDLINE=68091045; PubMed=4169272;
RA Fidalgo B.V., Najjar V.A.;
RT "The physiological role of the lymphoid system. VI. The stimulatory
RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic
RT activity of human polymorphonuclear leucocyte.";
RL Biochemistry 6:3386-3392(1967).
CC -!- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE
CC CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE
CC MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN.
CC TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC
CC ACTIVITY OF NEUTROPHILS.
DR PIR; A02147; A02147.
DR MIM; 191150; -.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006909; P:phagocytosis; NAS.
SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;

Query Match 40.0%; Score 8; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 EP 4
              :|
Db          2 KP 3
```

RESULT 2
THYL_PIG
ID THYL_PIG STANDARD; PRT; 3 AA.
AC P01151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
OS Sus scrofa (Pig),
OS Ovis aries (Sheep),
OS Bombina orientalis (Oriental fire-bellied toad), and
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823, 9940, 8346, 8316;
RN [1]
RP SEQUENCE.
RC SPECIES=Pig; TISSUE=Hypothalamus;
RX MEDLINE=70136150; PubMed=4984938;
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
RT "Structure of porcine thyrotropin releasing hormone.";
RL Biochemistry 9:1103-1106(1970).
RN [2]
RP SYNTHESIS.
RC SPECIES=Pig;
RX MEDLINE=70039904; PubMed=4982117;
RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
RT "The identity of chemical and hormonal properties of the thyrotropin
RT releasing hormone and pyroglutamyl-histidyl-proline amide.";
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
RN [3]
RP SEQUENCE.
RC SPECIES=Sheep; TISSUE=Hypothalamus;
RA Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R.,
RA Ward D.N.;
RT "The elucidation of the primary structure of the hypothalamic thyroid

```
RT   stimulating hormone releasing factor of ovine origin by means of mass
RT   spectrometry.";
RL   Org. Mass Spectrom. 5:221-228(1971).
RN   [4]
RP   SYNTHESIS.
RC   SPECIES=Sheep;
RX   MEDLINE=70163386; PubMed=4985794;
RA   Burgus R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RA   Guillemin R.;
RT   "Characterization of ovine hypothalamic hypophysiotropic
RT   TSH-releasing factor.";
RL   Nature 226:321-325(1970).
RN   [5]
RP   SEQUENCE.
RC   SPECIES=B.orientalis; TISSUE=Skin;
RX   MEDLINE=76138399; PubMed=815011;
RA   Yasuhara T., Nakajima T.;
RT   "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RL   Chem. Pharm. Bull. 23:3301-3303(1975).
RN   [6]
RP   SEQUENCE.
RC   SPECIES=N.viridescens;
RX   MEDLINE=75035605; PubMed=4214528;
RA   Grimm-Joergensen Y., McKelvy J.F.;
RT   "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT   viridescens) brain in vitro. Isolation and characterization of
RT   thyrotropin releasing factor.";
RL   J. Neurochem. 23:471-478(1974).
CC   -!- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
CC       IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
CC       NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
DR   PIR; A90919; RHTDTO.
DR   PIR; A92971; A92971.
DR   PIR; A93750; RHSHT.
KW   Amidation; Pyrrolidone carboxylic acid.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       3       3       AMIDATION.
SQ   SEQUENCE   3 AA;  380 MW;  7761F6B000000000 CRC64;

  Query Match             35.0%;  Score 7;  DB 1;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 P 4
              |
Db          3 P 3


RESULT 3
DCML_PSECH
ID   DCML_PSECH     STANDARD;      PRT;     4 AA.
AC   P19916;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE   dehydrogenase subunit L) (CO-DH L) (Fragment).
GN   CUTL.
OS   Pseudomonas carboxydohydrogena.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae.
OX   NCBI_TaxID=290;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90055678; PubMed=2818128;
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC       dioxide.
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC       acceptor.
CC   -!- COFACTOR: Molybdenum (molybdopterin).
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0140; PL0140.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER       4       4
SQ   SEQUENCE   4 AA;  441 MW;  7761E876F0000000 CRC64;

  Query Match             35.0%;  Score 7;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 P 4
              |
Db          4 P 4


RESULT 4
RM01_YEAST
ID   RM01_YEAST     STANDARD;      PRT;     4 AA.
AC   P36515;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   Mitochondrial 60S ribosomal protein L1 (YmL1) (Fragment).
GN   MRPL1.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=91285106; PubMed=2060626;
RA   Grohmann L., Graack H.-R., Kruft V., Choli T., Goldschmidt-Reisin S.,
RA   Kitakawa M.;
RT   "Extended N-terminal sequencing of proteins of the large ribosomal
RT   subunit from yeast mitochondria.";
RL   FEBS Lett. 284:51-56(1991).
DR   PIR; S17255; S17255.
DR   SGD; L0002681; MRPL1.
KW   Ribosomal protein; Mitochondrion.
FT   NON_TER       4       4
SQ   SEQUENCE   4 AA;  402 MW;  7771B2D5D0000000 CRC64;

  Query Match             35.0%;  Score 7;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 P 4
              |
Db          4 P 4


RESULT 5
EOSI_HUMAN
ID   EOSI_HUMAN     STANDARD;      PRT;     4 AA.
AC   P02731;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   21-JUL-1986 (Rel. 01, Last annotation update)
DE   Eosinophilotactic peptides.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=76078412; PubMed=1060093;
RA   Goetzl E.J., Austen K.F.;
RT   "Purification and synthesis of eosinophilotactic tetrapeptides of
RT   human lung tissue: identification as eosinophil chemotactic factor of
RT   anaphylaxis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
```

```
CC   -!- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
CC       (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
CC       (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
CC       EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
CC       OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
DR   GO; GO:0030105; P:anaphylaxis; IDA.
DR   GO; GO:0006935; P:chemotaxis; IDA.
FT   VARIANT       1      1       V -> A (IN OTHER PEPTIDE).
FT                                /FTId=VAR_005201.
SQ   SEQUENCE   4 AA;  390 MW;  6B05B862A0000000 CRC64;

  Query Match             25.0%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 E 3
              |
Db          4 E 4


RESULT 6
FAR3_HIRME
ID   FAR3_HIRME     STANDARD;      PRT;     4 AA.
AC   P42562;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide YLRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  598 MW;  69D4073B30000000 CRC64;

  Query Match             20.0%;  Score 4;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 L 2
              |
Db          2 L 2


RESULT 7
FLRF_HIRME
ID   FLRF_HIRME     STANDARD;      PRT;     4 AA.
AC   P42561;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FLRFamide.
OS   Hirudo medicinalis (Medicinal leech), and
OS   Helisoma trivolvis (Snail).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421, 27815;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=H.medicinalis;
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=H.trivolvis; TISSUE=Kidney;
RX   MEDLINE=94286417; PubMed=7912428;
RA   Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT   "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT   trivolvis.";
RL   Peptides 15:31-36(1994).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  582 MW;  69D40729A0000000 CRC64;

  Query Match             20.0%;  Score 4;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 L 2
              |
Db          2 L 2


RESULT 8
FLRN_ANTEL
ID   FLRN_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58707;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Antho-RNamide.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RX   MEDLINE=90319122; PubMed=1973541;
RA   Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA   Reinscheid R.K., Nothacker H.-P., Staley A.L.;
RT   "Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
RT   anemone neuropeptide containing an unusual amino-terminal blocking
RT   group.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron-specific.
CC   -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.
DR   PIR; A35779; A35779.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       L-3-PHENYLLACTYL.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  549 MW;  64540729A0000000 CRC64;

  Query Match             20.0%;  Score 4;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 L 2
              |
Db          2 L 2


RESULT 9
LUXE_VIBFI
ID   LUXE_VIBFI     STANDARD;      PRT;     3 AA.
AC   P24272;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-
DE   protein synthetase) (Fragment).
GN   LUXE.
OS   Vibrio fischeri.
```

```
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=668;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91072226; PubMed=2254256;
RA   Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;
RT   "A new Vibrio fischeri lux gene precedes a bidirectional termination
RT   site for the lux operon.";
RL   J. Bacteriol. 172:6797-6802(1990).
CC   -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC       IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
CC       FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
CC       SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC   -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
CC       an acyl-protein thiolester.
CC   -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62812; -; NOT_ANNOTATED_CDS.
KW   Luminescence; Ligase.
FT   NON_TER       1       1
SQ   SEQUENCE   3 AA;  374 MW;  6AA3303000000000 CRC64;

  Query Match             15.0%;  Score 3;  DB 1;  Length 3;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 V 1
              :
Db          1 I 1


RESULT 10
FYRI_ANTEL
ID   FYRI_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58706;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Antho-RIamide I [Contains: Antho-RIamide II].
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92270459; PubMed=1821096;
RA   Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA   Grimmelikhuijzen C.J.P.;
RT   "Isolation of two novel neuropeptides from sea anemones: the unusual,
RT   biologically active L-3-phenyllactyl-Tyr-Arg-Ile-NH2 and its .
RT   des-phenyllactyl fragment Tyr-Arg-Ile-NH2.";
RL   Peptides 12:1165-1173(1991).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=93391436; PubMed=8397415;
RA   McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT   "The expansion behaviour of sea anemones may be coordinated by two
RT   inhibitory neuropeptides, Antho-KAamide and Antho-RIamide.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC   -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC       groups. May be involved in the expansion phase of feeding
CC       behaviour in sea anemones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron-specific.
KW   Neuropeptide; Amidation.
FT   CHAIN         1       4       ANTHO-RIAMIDE I.
FT   CHAIN         2       4       ANTHO-RIAMIDE II.
FT   MOD_RES       1       1       L-3-PHENYLLACTYL.
FT   MOD_RES       4       4       AMIDATION.
SQ   SEQUENCE   4 AA;  598 MW;  60441B59A0000000 CRC64;

  Query Match             15.0%;  Score 3;  DB 1;  Length 4;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 V 1
              :
Db          4 I 4


RESULT 11
ACH1_ACHFU
ID   ACH1_ACHFU     STANDARD;      PRT;     4 AA.
AC   P35904;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Achatin-I.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RC   STRAIN=Ferussac; TISSUE=Ganglion;
RX   MEDLINE=89273551; PubMed=2597281;
RA   Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
RA   Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
RA   Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
RT   "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
RT   fulica Ferussac containing a D-amino acid residue.";
RL   Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
RN   [2]
RP   CHARACTERIZATION.
RC   STRAIN=Ferussac; TISSUE=Heart atrium;
RX   MEDLINE=91264856; PubMed=1675568;
RA   Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
RA   Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
RT   "Purification of achatin-I from the atria of the African giant snail,
RT   Achatina fulica, and its possible function.";
RL   Biochem. Biophys. Res. Commun. 177:847-853(1991).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY.
RX   MEDLINE=93014529; PubMed=1399265;
RA   Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
RA   Iwashita T., Nomoto K.;
RT   "Crystal structure and molecular conformation of achatin-I
RT   (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
RT   D-amino acid residue.";
RL   Int. J. Pept. Protein Res. 39:258-264(1992).
CC   -!- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
CC       AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
CC       NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
CC       HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
DR   PIR; A32480; A32480.
KW   Hormone; D-amino acid.
FT   MOD_RES       2       2       D-PHENYLALANINE.
SQ   SEQUENCE   4 AA;  408 MW;  6AADD9C810000000 CRC64;

  Query Match             10.0%;  Score 2;  DB 1;  Length 4;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 E 3
              :
Db          4 D 4
```

```
RESULT 12
DCMS_PSECH
ID   DCMS_PSECH     STANDARD;      PRT;     4 AA.
AC   P19918;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO
DE   dehydrogenase subunit S) (CO-DH S) (Fragment).
GN   CUTS.
OS   Pseudomonas carboxydohydrogena.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae.
OX   NCBI_TaxID=290;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90055678; PubMed=2818128;
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC       dioxide.
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC       acceptor.
CC   -!- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0146; PL0146.
KW   Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
FT   NON_TER       4       4
SQ   SEQUENCE   4 AA;  420 MW;  6DD33DD6F0000000 CRC64;

  Query Match             10.0%;  Score 2;  DB 1;  Length 4;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 L 2
               :
Db          1 M 1


RESULT 13
FAR4_HIRME
ID   FAR4_HIRME     STANDARD;      PRT;     4 AA.
AC   P42563;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide YMRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4       4       AMIDATION.
SQ   SEQUENCE   4 AA;  616 MW;  69D4068B30000000 CRC64;

  Query Match             10.0%;  Score 2;  DB 1;  Length 4;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 L 2
               :
Db          2 M 2


RESULT 14
FMRF_MACNI
ID   FMRF_MACNI     STANDARD;      PRT;     4 AA.
AC   P01162;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   FMRFamide (Peak C) (Cardioexcitatory neuropeptide).
OS   Macrocallista nimbosa (Sun-ray clam),
OS   Nereis virens (Sandworm),
OS   Hirudo medicinalis (Medicinal leech), and
OS   Helisoma trivolvis (Snail).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida;
OC   Veneroidea; Veneridae; Macrocallista.
OX   NCBI_TaxID=6594, 6353, 6421, 27815;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;
RX   MEDLINE=77215956; PubMed=877582;
RA   Price D.A., Greenberg M.J.;
RT   "Structure of a molluscan cardioexcitatory neuropeptide.";
RL   Science 197:670-671(1977).
RN   [2]
RP   SEQUENCE, AND CHARACTERIZATION.
RC   SPECIES=M.nimbosa; TISSUE=Ganglion;
RX   MEDLINE=78012038; PubMed=909875;
RA   Price D.A., Greenberg M.J.;
RT   "Purification and characterization of a cardioexcitatory neuropeptide
RT   from the central ganglia of a bivalve mollusc.";
RL   Prep. Biochem. 7:261-281(1977).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=N.virens;
RX   MEDLINE=90259866; PubMed=2342992;
RA   Krajniak K.G., Price D.A.;
RT   "Authentic FMRFamide is present in the polychaete Nereis virens.";
RL   Peptides 11:75-77(1990).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=H.medicinalis;
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
RN   [5]
RP   SEQUENCE.
RC   SPECIES=H.trivolvis; TISSUE=Kidney;
RX   MEDLINE=94286417; PubMed=7912428;
RA   Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT   "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT   trivolvis.";
RL   Peptides 15:31-36(1994).
CC   -!- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL
CC       ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULARIZATION OF
CC       CARDIAC CONTRACTION.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; A01426; ECNK.
DR   PIR; A60418; A60418.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4       4       AMIDATION.
SQ   SEQUENCE   4 AA;  600 MW;  69D40699A0000000 CRC64;

  Query Match             10.0%;  Score 2;  DB 1;  Length 4;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 L 2
               :
Db          2 M 2
```

```
RESULT 15
OCP1_OCTMI
ID   OCP1_OCTMI     STANDARD;      PRT;     4 AA.
AC   P58648;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cardioactive peptides Ocp-1/Ocp-2.
OS   Octopus minor (Octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC   Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=89766;
RN   [1]
RP   SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC   TISSUE=Brain;
RX   MEDLINE=20336815; PubMed=10876044;
RA   Iwakoshi E., Hisada M., Minakata H.;
RT   "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT   Octopus minor.";
RL   Peptides 21:623-630(2000).
CC   -!- FUNCTION: Cardioactive; has both positive chronotropic and
CC       inotropic effects on the heart. Ocp-2 is a 1000 time less
CC       active than Ocp-1.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Ocp-2 has L-Phe instead of D-Phe.
CC   -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW   Hormone; D-amino acid.
FT   MOD_RES       2      2       D-PHENYLALANINE.
SQ   SEQUENCE   4 AA;  394 MW;  6AA879C810000000 CRC64;

  Query Match             10.0%;  Score 2;  DB 1;  Length 4;
  Best Local Similarity   0.0%;  Pred. No. 1.3e+05;
  Matches    0;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 E 3
              :
Db          4 D 4


Search completed: November 25, 2003, 14:07:03
Job time : 12 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:04:37 ; Search time 34 Seconds
(without alignments)
30.359 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9	45.0	4	11	Q08433	Q08433 rattus sp.

ALIGNMENTS

RESULT 1
Q08433
ID Q08433 PRELIMINARY; PRT; 4 AA.
AC Q08433;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Bilirubin UDP-glucuronosyltransferase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gunn;
RX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
DR EMBL; S38636; AAB19259.1; -.
KW Transferase.
FT NON_TER 1 1
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 45.0%; Score 9; DB 11; Length 4;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VLE 3
              ||:
Db          2 VLK 4
```

Search completed: November 25, 2003, 14:07:49
Job time : 35 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:00:57 ; Search time 41 Seconds
(without alignments)
15.486 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 13293

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	4	21	AAB28663	Chemotactic peptid
2	20	100.0	4	22	AAB29558	Adhesion-modulator
3	20	100.0	4	24	ABP58757	Macrophage recruit
4	16	80.0	4	20	AAY03962	Interleukin-1 beta
5	14	70.0	4	20	AAY15640	Peptide used to ma
6	14	70.0	4	21	AAB28671	Chemotactic peptid
7	14	70.0	4	21	AAB22909	Granzyme B substra
8	14	70.0	4	21	AAY79611	Granzyme-B substra
9	14	70.0	4	21	AAY80804	Fluorophore-labell
10	14	70.0	4	22	AAU79766	Chicken bone deriv
11	14	70.0	4	22	AAB82937	Granzyme B peptide
12	14	70.0	4	22	AAG62561	Cresyl violet subs
13	14	70.0	4	22	AAB29759	Escherichia coli r
14	14	70.0	4	23	AAE28993	Human thrombin cle
15	14	70.0	4	23	ABG94471	Protease biosensor
16	14	70.0	4	23	AAE19128	Granzyme B peptide
17	14	70.0	4	23	AAM50406	Tumour associated
18	14	70.0	4	23	AAM50407	Tumour associated
19	14	70.0	4	23	AAM50466	Tumour associated
20	14	70.0	4	23	AAM50470	Tumour associated
21	14	70.0	4	23	AAM50471	Tumour associated
22	13	65.0	4	8	AAP71434	Immunomodulator pe
23	13	65.0	4	13	AAR23726	ACE inhibitor pept
24	13	65.0	4	16	AAW93983	Antineoplastic pep
25	13	65.0	4	19	AAW62052	Human erythropoiet
26	13	65.0	4	20	AAY15644	Peptide used to ma
27	13	65.0	4	20	AAW83445	Human growth hormo
28	13	65.0	4	21	AAY80808	Fluorophore-labell
29	13	65.0	4	21	AAY61373	Cadherin-7 cell ad
30	13	65.0	4	22	AAU79767	Chicken bone deriv
31	13	65.0	4	22	AAB30764	Peptide which is u
32	13	65.0	4	23	ABG61371	Tick Ixolaris firs
33	12	60.0	2	22	AAB91036	Thyrotropin releas
34	12	60.0	2	24	ABU06374	Maize Starch synth
35	12	60.0	3	16	AAR83188	Tripeptide having
36	12	60.0	3	19	AAW56205	Anti-inflammatory
37	12	60.0	3	21	AAY56894	Peptide used as an
38	12	60.0	4	5	AAP40420	Animal growth prom
39	12	60.0	4	5	AAP40424	Animal growth prom
40	12	60.0	4	9	AAP82910	Acetylcholinestera
41	12	60.0	4	9	AAP80424	Sequence of peptid
42	12	60.0	4	14	AAR38100	Diuretic hormone b
43	12	60.0	4	14	AAR38441	Diuretic hormone b
44	12	60.0	4	14	AAR42094	Diuretic hormone d
45	12	60.0	4	14	AAR42095	Diuretic hormone d

ALIGNMENTS

RESULT 1
AAB28663
ID AAB28663 standard; peptide; 4 AA.
XX
AC AAB28663;
XX
DT 13-FEB-2001 (first entry)
XX
DE Chemotactic peptide pepJ'.
XX
KW Chemotactic; osteopontin; vulnerary; antiarthritic; antipsoriatic;
KW cytostatic; antitumour; antiinflammatory; osteopathic;
KW wound healing; cell migration; chemotaxis; atherosclerosis; cancer;
KW angiogenic-associated disease; arthritis; psoriasis; haemangioma;
KW ocular neovascularisation; cell apoptosis; nitrous oxide production;
KW inflammation; osteoporosis; immune disease.
XX
OS Mammalia.
OS Synthetic.
XX
PN WO200063247-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-US10344.
XX
PR 15-APR-1999; 99US-0129764.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX

DR WPI; 2000-687159/67.
XX
PT New osteopontin-derived chemotactic and inhibitory peptides, useful for
PT promoting scarless wound healing, modulating cellular chemotaxis,
PT treating formation of atherosclerotic plaques and preventing metastasis
PT -
XX
PS Claim 17; Page 43; 54pp; English.
XX
CC The present sequence is an osteopontin-derived chemotactic peptide.
CC Such chemotactic peptides are useful for promoting scarless wound
CC healing, modulating chemotaxis and promoting cell migration to a target
CC site in a cell of a subject. They are also used for modulating cellular
CC chemotaxis in a mammalian cell such as smooth muscle cell, a macrophage,
CC an endothelial cell, a vascular cell and a tumourigenic cell. They are
CC useful for treating the formation of atherosclerotic plaques in a
CC subject. The peptides are used for preventing metastasis, treating an
CC angiogenic-associated disease such as arthritis, psoriasis, haemangioma,
CC tumour metastasis or ocular neovascularisation. They are also used for
CC activating cell apoptosis, for modulating nitrous oxide production and
CC for inducing chemotaxis. The peptides are useful for diagnosing, treating
CC and preventing tumour metastasis, inflammation, osteoporosis and immune
CC diseases. They can also be used to enhance an immune response by
CC attracting macrophages.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 2
AAB29558
ID AAB29558 standard; peptide; 4 AA.
XX
AC AAB29558;
XX
DT 14-FEB-2001 (first entry)
XX
DE Adhesion-modulatory peptide, SEQ ID NO:15.
XX
KW Adhesion modulatory peptide; target cell adhesion; cell adhesion;
KW endothelial cell; fibroblast; macrophage; neutrophil; myofibroblast;
KW collagen; glycosaminoglycan; extracellular matrix; synthetic substrate;
KW vascular growth; wound healing; keloid formation; scarring; fibrosis;
KW anticoagulant; vulnerary; immunomodulatory; antibacterial; anticancer;
KW antitumorigenic; anti-CD44 activity; prosthesis; implant.
XX
OS Synthetic.
XX
PN WO200063236-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-US10329.
XX
PR 16-APR-1999; 99US-0129709.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX
DR WPI; 2001-007002/01.
XX
PT Novel adhesive modulatory peptides useful for modulating adhesion of
PT target cells such as endothelial cells, fibroblasts, macrophages to
PT substrate such as polyvinyl surfaces, collagen -
XX
PS Claim 8; Page 27; 35pp; English.
XX
CC The invention relates to novel adhesion modulatory peptides (AAB29544-
CC AAB29558) which modulate the adhesion of a target cell to a substrate.
CC The invention also encompasses substrates and devices treated with a
CC peptide of the invention; compositions comprising a peptide of the
CC invention for in vivo use; and analogues, fragments and chemical
CC derivatives of the peptides of the invention. The peptides are useful for
CC modulating the adhesion to a substrate of target cells such as
CC endothelial cells, fibroblasts, macrophages, neutrophils or
CC myofibroblasts. The substrate may be a substrate that is found in the
CC body of a patient, e.g., collagen or hyaluronic acid, or may be a
CC synthetic substrate e.g., a polyvinyl surface, titanium or PGA. The
CC peptides are useful for regulating vessel growth during wound healing
CC and/or in the treatment of damage resulting from vascular disease; for
CC inhibiting or preventing cellular apoptosis; in the treatment of
CC fibrosis, in particular in the clearing of debris; to minimise wound
CC contraction, thereby reducing keloid tissue formation and scarring; and
CC as anti clotting agents. The peptides also have an immunomodulatory
CC effect, and an antibacterial effect by adhering to neutrophils.
CC Additionally, peptides of the invention have an anticancer effect by
CC competing for alpha-v-beta3 integrin binding on the cell surface, and an
CC antitumorigenic effect by having anti-CD44 activity. The peptides are
CC useful for stimulating and/or enhancing cell attachment to polymer
CC scaffolds, to enhance tissue growth and for coating medical devices,
CC including prostheses and implants (e.g., vascular implants). The present
CC sequence represents a specifically claimed adhesion modulatory peptide of
CC the invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 22; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 3
ABP58757
ID ABP58757 standard; peptide; 4 AA.
XX
AC ABP58757;
XX
DT 04-APR-2003 (first entry)
XX
DE Macrophage recruitment inhibitor peptide, SEQ ID NO:7.
XX
KW Minicell display; random peptide library; drug discovery; screening;
KW fusion protein; outer membrane protein;
KW Rickettsia rickettsii 17K antigen; bioactive peptide;
KW macrophage recruitment inhibitor; osteopontin; C5a; fibronectin.
XX
OS Synthetic.
XX
PN WO200272759-A2.
XX
PD 19-SEP-2002.
XX
PF 06-MAR-2002; 2002WO-US06921.
XX
PR 07-MAR-2001; 2001US-274039P.
PR 20-JUL-2001; 2001US-306946P.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX
DR WPI; 2003-046738/04.
XX
PT Minicell display method for screening peptide libraries for peptide

PT candidates that can bind and modulate a particular biological process,
PT comprises identifying minicell hosts bound to the binding partner -
XX
PS Claim 22; Page 36; 66pp; English.
XX
CC The invention relates to minicell display methods for the generation
CC and screening of random peptide libraries for peptide candidates which
CC can bind to and modulate the activity of target molecules involved in
CC biological processes. The methods involve expressing a peptide library as
CC a fusion protein with a bacterial outer membrane protein (preferably the
CC 17K antigen of Rickettsia rickettsii) in an anucleate bacterially-derived
CC minicell host, and then contacting the minicell hosts with a target
CC molecule. Minicells comprising a library peptide which has bound to the
CC target molecule are separated from the unbound minicells, and the
CC peptides of interest analysed. The invention also encompasses methods for
CC making a minicell DNA library, increasing the diversity of a minicell DNA
CC library by in vivo mutagenesis, screening a minicell DNA library, and
CC purifying minicells. The methods of the invention are useful for
CC identifying minicell hosts bound to a target molecule and may be used in
CC drug discovery. The method increases the size of the peptides to be
CC screened and the diversification of the library to be screened, thereby
CC increasing the number of potential peptides that can modulate a
CC particular biological response or mechanism. Sequences ABP58757-ABP58803
CC represent bioactive peptides identified using the minicell display
CC method of the invention. The present sequence was characterised as
CC an inhibitor of macrophage recruitment by osteopontin, C5a and
CC fibronectin.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 24; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 4
AAY03962
ID AAY03962 standard; peptide; 4 AA.
XX
AC AAY03962;
XX
DT 23-JUN-1999 (first entry)
XX
DE Interleukin-1 beta converting enzyme binding moiety.
XX
KW Cysteine protease inhibitor; oxadiazole; thiadiazole; 1,2,4-triazole;
KW interleukin-1 beta converting enzyme inhibitor; IL-1; ICE; antitumour;
KW anticancer; antimicrobial; antibacterial; antiviral; antiallergic;
KW antiinflammatory; herbicide; fungicide; pesticide; multiple sclerosis;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW ischaemic-reperfusion injury.
XX
OS Synthetic.
XX
PN WO9849190-A2.
XX
PD 05-NOV-1998.
XX
PF 24-APR-1998; 98WO-US08259.
XX
PR 23-APR-1998; 98US-0044819.
PR 25-APR-1997; 97US-0044819.
XX
PA (CORT-) CORTECH INC.
XX
PI Cheronis JC, Gerrity JI, Goodfellow VS, Gyorkos AC;
PI Leimer AH, Spruce LW, Young JM;
XX
DR WPI; 1999-131683/11.
XX
PT New substituted oxadiazole cysteine protease inhibitors - useful
PT e.g. for inhibiting cancer cell growth, microbial cell or viral
PT growth or treating inflammatory and degenerative diseases
XX
PS Claims 61, 62; Pages 67, 68; 82pp; English.
XX
CC New compounds are presented which consist of a cysteine protease
CC binding moiety attached by its C-terminal to an oxadiazole, thiadiazole
CC or 1,2,4-triazole moiety. These compounds are cysteine protease
CC inhibitors useful for inhibiting the enzymatic activity of calpain,
CC cathepsin, caspase (e.g. human interleukin-1 beta converting enzyme),
CC viral or microbial cysteine protease (e.g. human coronavirus or
CC gingipain), or protozoan cysteine protease (e.g. Trypanosoma,
CC Schistosoma, Leishmania or Plasmodium protease). They can be used
CC e.g. for: inhibiting cancer cell growth, tumour progression or tumour
CC metastasis, or microbial cell or viral growth (e.g. inhibiting
CC hepatitis A virus 3C proteinase, hepatitis C virus endopeptidase 2,
CC picornain 3C rhinovirus protease, foot and mouth disease virus L
CC proteinase, encephalomyelitis virus endopeptidase 2, picornain 2A
CC protease); treating allergic response symptoms (e.g. inhibiting
CC protease Der pI); treating neurodegenerative disorders (e.g.
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis) and
CC disorders resulting from ischaemic-reperfusion injury (e.g. stroke,
CC myocardial infarction, transplantation, vascular injury or
CC cardiovascular collapse or shock); treating inflammatory and
CC degenerative diseases (e.g. rheumatoid arthritis, osteoarthritis or
CC periodontal disease); or treating pulmonary diseases (e.g. asthma or
CC emphysema). They can also be used for detecting or quantifying cysteine
CC protease activity in a sample, mixture, biological fluid or tissue; for
CC purifying cysteine protease in a sample; or as antibacterial agents,
CC herbicides, fungicides or pesticides.
XX
SQ Sequence 4 AA;

Query Match 80.0%; Score 16; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
|||
Db 1 LEP 3

RESULT 5
AAY15640
ID AAY15640 standard; Peptide; 4 AA.
XX
AC AAY15640;
XX
DT 27-JUL-1999 (first entry)
XX
DE Peptide used to make fluorescent reporter molecules.
XX
KW Fluorogenic; fluorescent reporter molecule; enzyme substrate;
KW apoptosis; protease; peptidase; apoptosis cascade; cancer;
KW chemotherapeutic agent; cell death; viral protease activity.
XX
OS Synthetic.
XX
PN WO9918856-A1.
XX
PD 22-APR-1999.
XX
PF 09-OCT-1998; 98WO-US21231.
XX
PR 03-MAR-1998; 98US-0033661.
PR 10-OCT-1997; 97US-0061582.
XX
PA (CYTO-) CYTOVIA INC.
XX
PI Cai SX, Drewe JA, Keana JFW, Weber E, Zhang H;
XX

DR WPI; 1999-312448/26.
XX
PT New fluorogenic or fluorescent reporter molecules
XX
PS Claim 6; Page 165; 202pp; English.
XX
CC AAY15618-Y15759 represent peptides used to make the fluorogenic or
CC fluorescent reporter molecules of the invention. These molecules
CC contain a peptide moiety (e.g. present sequence) which acts as a
CC substrate for enzymes involved in apoptosis or protease or peptidase
CC enzymes. The compounds can be used as fluorogenic or fluorescent
CC substrates for enzymes. Depending on the peptide moiety used, the
CC fluorescent molecules can be used for detecting or measuring the
CC activity of an enzyme involved in the apoptosis cascade in cells; to
CC determine whether a test compound has an effect on an enzyme involved
CC in the apoptosis cascade in cells; for determining the sensitivity of
CC an animal with cancer to treatment with chemotherapeutic agents or
CC determining whether a test substance inhibits, prevents, causes or
CC enhances cell death of test cells; for detecting or measuring the
CC activity of a viral protease in cells; for determining whether a test
CC compound has an effect on the activity of a viral protease in cells;
CC and for measuring the activity or determining whether a test substance
CC has an effect on the activity of a protease or peptidase in cells.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 20; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
:||
Db 1 IEP 3

RESULT 6
AAB28671
ID AAB28671 standard; peptide; 4 AA.
XX
AC AAB28671;
XX
DT 13-FEB-2001 (first entry)
XX
DE Chemotactic peptide pepL.
XX
KW Chemotactic; osteopontin; vulnerary; antiarthritic; antipsoriatic;
KW cytostatic; antitumour; antiinflammatory; osteopathic;
KW wound healing; cell migration; chemotaxis; atherosclerosis; cancer;
KW angiogenic-associated disease; arthritis; psoriasis; haemangioma;
KW ocular neovascularisation; cell apoptosis; nitrous oxide production;
KW inflammation; osteoporosis; immune disease.
XX
OS Mammalia.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
XX
PN WO200063247-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-US10344.
XX
PR 15-APR-1999; 99US-0129764.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Ashkar S;
XX
DR WPI; 2000-687159/67.
XX
PT New osteopontin-derived chemotactic and inhibitory peptides, useful for
PT promoting scarless wound healing, modulating cellular chemotaxis,
PT treating formation of atherosclerotic plaques and preventing metastasis
PT -
XX
PS Claim 17; Page 43; 54pp; English.
XX
CC The present sequence is an osteopontin-derived chemotactic peptide.
CC Such chemotactic peptides are useful for promoting scarless wound
CC healing, modulating chemotaxis and promoting cell migration to a target
CC site in a cell of a subject. They are also used for modulating cellular
CC chemotaxis in a mammalian cell such as smooth muscle cell, a macrophage,
CC an endothelial cell, a vascular cell and a tumourigenic cell. They are
CC useful for treating the formation of atherosclerotic plaques in a
CC subject. The peptides are used for preventing metastasis, treating an
CC angiogenic-associated disease such as arthritis, psoriasis, haemangioma,
CC tumour metastasis or ocular neovascularisation. They are also used for
CC activating cell apoptosis, for modulating nitrous oxide production and
CC for inducing chemotaxis. The peptides are useful for diagnosing, treating
CC and preventing tumour metastasis, inflammation, osteoporosis and immune
CC diseases. They can also be used to enhance an immune response by
CC attracting macrophages.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 21; Length 4;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
:|:|
Db 1 MLDP 4

RESULT 7
AAB22909
ID AAB22909 standard; peptide; 4 AA.
XX
AC AAB22909;
XX
DT 10-JAN-2001 (first entry)
XX
DE Granzyme B substrate recognition sequence, SEQ ID NO:100.
XX
KW Biodetector protein; fusion protein; recognition site;
KW cellular targetting sequence; cellular localisation; fluorescent protein;
KW protease activity detection; toxin detection; cellular stress detection;
KW drug discovery; cell based screening; protease recognition site;
KW cleavage site.
XX
OS Unidentified.
XX
PN WO200050872-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04794.
XX
PR 26-FEB-1999; 99US-0122152.
PR 08-MAR-1999; 99US-0123399.
PR 12-JUL-1999; 99US-0352171.
XX
PA (CELL-) CELLOMICS INC.
XX
PI Giuliano KA, Kapur R;
XX
DR WPI; 2000-594086/56.
DR N-PSDB; AAA93400.
XX
PT Automated cell-based characterization of toxin by contacting cells
PT containing luminescent reporter molecules with test substance and
PT analyzing optically -
XX

PS Example 11; Fig 29B; 336pp; English.
XX
CC The invention relates to systems, methods and reagents for cell-based
CC screening or detection of compounds which affect particular biological
CC functions. The methods of the invention utilise fluorescent biodetector
CC molecules which, when acted on by a compound of interest, cause an
CC alteration in the cellular distribution of at least the fluorescent
CC moiety. In one embodiment, the biosensors comprise heat shock proteins
CC (HSPs) fused to a fluorescent protein (e.g., jellyfish green fluorescent
CC protein (GFP), or derivatives thereof). Such biosensors are located in
CC the cytoplasm, but on stress activation translocate to the nucleus. In
CC another embodiment biodetector proteins can be used to detect protease
CC activity. Such protease biodetector fusion proteins comprise one or more
CC fluorescent proteins; a recognition signal which is cleaved by the
CC protease; and at least one cellular localisation signal. The latter two
CC components may be components of a single protein which is acted upon by
CC the protease, or may be from heterologous sources. Due to the
CC localisation signal, the biodetector protein is localised to a
CC particular region of the cell. Once acted on by the protease of interest,
CC the fluorescent protein is cleaved from the localisation sequence, and
CC is free to migrate to other locations within the cell. The presence of a
CC second localisation signal attached to the fluorescent protein enables
CC the fluorescent protein to be directed to a different cellular
CC compartment after cleavage of the protease recognition sequence. The
CC change in distribution of the fluorescent protein can be detected using
CC imaging methods with a high degree of spatial resolution. The methods
CC and biosensors of the invention can be used to investigate a wide range
CC of cellular activities and to screen compounds which modulate these
CC activities. Biosensors containing a recognition site for caspase, for
CC example, may be used for the screening of compounds which modulate
CC apoptosis, while biosensors containing other protease recognition sites
CC may be used for the detection of proteolytic toxins (such as anthrax
CC lethal factor). The method provides improved target validation and
CC candidate compound optimisation by combining many cell screening formats
CC with fluorescence-based molecular reagents and computer-based feature
CC extraction, data analysis and automation, resulting in increased
CC quantity and speed of data collection and faster evaluation of drug
CC candidates. Sequences AAB22886-B22920 and AAB22935 represent protease
CC recognition sites which may used as components of biosensor fusion
CC proteins of the invention.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 21; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
:||
Db 1 IEP 3

RESULT 8
AAY79611
ID AAY79611 standard; Peptide; 4 AA.
XX
AC AAY79611;
XX
DT 29-AUG-2000 (first entry)
XX
DE Granzyme-B substrate recognition sequence.
XX
KW Protease; biosensor; granzyme-B; substrate recognition sequence;
KW cell screening; assay; analysis; drug discovery.
XX
OS Unidentified.
XX
PN WO200026408-A2.
XX
PD 11-MAY-2000.
XX
PF 29-OCT-1999; 99WO-US25431.
XX
PR 30-OCT-1998; 98US-0106308.
PR 26-MAY-1999; 99US-0136078.
XX
PA (CELL-) CELLOMICS INC.
XX
PI Guiliano KA, Bright G, Olson K, Burroughs-Tencza S;
XX
DR WPI; 2000-365644/31.
DR N-PSDB; AAA27600.
XX
PT Recombinant nucleic acid encoding a protease biosensor useful for
PT fluorescence based cell and molecular biochemical assays for drug
PT discovery comprising three operably linked nucleic acid sequences -
XX
PS Claim 15; Fig 29B; 218pp; English.
XX
CC The present sequence is that of a granzyme-B substrate
CC recognition sequence, which can be included in novel recombinant
CC protease biosensors (PBs) of the invention. The PBs (see AAY79638-54)
CC comprise: a first domain (see AAY79579-87) comprising at least 1
CC detectable polypeptide signal; a second domain (see AAY79588-622)
CC comprising at least 1 protease recognition site, such as the
CC present sequence; and a third domain (see AAY79623-37) comprising at
CC least 1 reactant target sequence. A recombinant nucleic acid (see
CC AAA27627-43) encoding the PB, an expression vector, and a genetically
CC engineered host cell are also claimed. A claimed method for
CC identifying compounds that modify protease activity in a cell
CC involves contacting a host cell that possesses the recombinant PB
CC with a test compound, and determining the PB distribution in the
CC host cell, where changes in the distribution of the PB are
CC correlated with modification of protease activity by the test
CC compound. Claimed kits for identifying compounds that modify
CC protease activity in a host cell include the recombinant nucleic
CC acid, or the recombinant PB, or the vector, or the host cell. The
CC PB is useful in high content screens to detect in vivo activation
CC of enzymatic activity, and to identify specific activity based on
CC cleavage of a known recognition motif.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 21; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
:||
Db 1 IEP 3

RESULT 9
AAY80804
ID AAY80804 standard; peptide; 4 AA.
XX
AC AAY80804;
XX
DT 22-MAY-2000 (first entry)
XX
DE Fluorophore-labelled granzyme B substrate peptide, SEQ ID NO:23.
XX
KW Protease substrate; fluorescent label; fluorophore; rhodamine;
KW blocking group; halobenzoyl group; cleavage; caspase; viral protease;
KW methionine aminopeptidase type 2; MetAP-2; drug screening.
XX
OS Synthetic.
XX
PN WO200004914-A1.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US16423.
XX
PR 21-JUL-1998; 98US-0093642.
XX

PA (CYTO-) CYTOVIA INC.
PA (ZHAN/) ZHANG H.
PA (CAIS/) CAI S X.
PA (DREW/) DREWE J A.
PA (YANG/) YANG W.
XX
PI Zhang H, Cai SX, Drewe JA, Yang W;
XX
DR WPI; 2000-195079/17.
XX
PT New fluorescently labeled amino acids or peptides, used as substrates
PT for detecting enzymes or their modulators, e.g. anticancer or antiviral
PT agents, contains a halobenzoyl N-blocking group -
XX
PS Claim 3; Page 100; 174pp; English.
XX
CC The invention relates to fluorescently labelled peptides containing
CC a halobenzoyl group on the fluorophore. They are of the structure
CC peptide-Y-Z, where Z represents a halo-substituted benzoyl blocking
CC group, Y is a fluorescent or fluorigenic moiety (preferably a
CC rhodamine), and the peptide-Y bond is cleavable by the enzyme being
CC assayed. The labelled peptides are reporters for detecting intracellular
CC proteolytic enzymes, particularly caspases and other enzymes involved in
CC apoptosis; viral proteases (e.g., HIV, herpes simplex virus-1, human
CC cytomegalovirus and hepatitis C virus proteases); and methionine
CC aminopeptidase type 2 (MetAP-2). The peptides are particularly used to
CC identify modulators of these enzymes which may be potentially useful as
CC agents for treating conditions such as cancer, neurodegeneration,
CC autoimmune diseases, myocardial infarction and virla infection.
CC Modulators identified may also be used to prolong the life of cells being
CC cultured for recombinant protein production, or to monitor the treatment
CC of cancer with chemotherapeutic agents. Inhibitors of MetAP-2 are
CC potential anti- angiogenic or anticancer agents. Sequences
CC AAY80782-Y80910 represent peptides, some of which are specifically
CC claimed, which may be used in assay methods according to the invention.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 21; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 LEP 4
              :||
Db          1 IEP 3
```

RESULT 10
AAU79766
ID AAU79766 standard; peptide; 4 AA.
XX
AC AAU79766;
XX
DT 30-JUL-2002 (first entry)
XX
DE Chicken bone derived peptide #1, useful as ACE inhibitor.
XX
KW Angiotensin converting enzyme inhibitor; ACEI; health food;
KW food additive; pharmaceutical product; lowering blood pressure;
KW hypertension; chicken essence by-product; chicken bone.
XX
OS Gallus sp.
XX
PN JP2001163896-A.
XX
PD 19-JUN-2001.
XX
PF 03-APR-2000; 2000JP-0105410.
XX
PR 01-DEC-1999; 99TW-0120997.
XX
PA (FOOD-) FOOD IND RES & DEV INST.
XX
DR WPI; 2001-574451/65.
XX
PT New peptides capable of inhibiting angiotensin converting enzyme for
PT use in pharmaceutical compositions, in health foods and as food
PT additives -
XX
PS Claim 1; Page 2; 39pp; Japanese.
XX
CC The present invention relates to novel peptides having angiotensin
CC converting enzyme (ACE) inhibitory activity. Also described are
CC methods for preparing products having ACE inhibitory activity, a
CC pharmaceutical composition for use in the inhibition of ACE, and
CC a health food or a food additive for use in the inhibition of
CC ACE. The peptides of the invention are useful as ACE inhibitors (ACEI).
CC They are useful in pharmaceutical products and health foods, and as
CC food additives. Such compositions maybe used for lowering blood pressure
CC and treating hypertension. The manufacturing cost of the peptides of the
CC invention is considerably less than prior art as the peptide can be
CC produced from waste materials. Also the values of chicken essence
CC by-products can be improved. The present peptide derived from chicken
CC bone is useful as an ACE inhibitor.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 22; Length 4;
Best Local Similarity 75.0%; Pred. No. 9.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 VLEP 4
              || |
Db          1 VLPP 4
```

RESULT 11
AAB82937
ID AAB82937 standard; Peptide; 4 AA.
XX
AC AAB82937;
XX
DT 21-DEC-2001 (first entry)
XX
DE Granzyme B peptide substrate.
XX
KW Granzyme B; caspase; prodrug; therapy; ADEPT; cancer; tumour;
KW inflammation; infection.
XX
OS Synthetic.
XX
PN WO200162300-A2.
XX
PD 30-AUG-2001.
XX
PF 22-FEB-2001; 2001WO-US05709.
XX
PR 24-FEB-2000; 2000US-184779P.
XX
PA (GETH) GENENTECH INC.
XX
PI Carter PJ, Gazzard L;
XX
DR WPI; 2001-611217/70.
XX
PT Delivering active agents to cell-types of interest for therapeutic
PT purposes, comprises administering a cell-type targeted caspase
PT conjugate and a pro-agent which is converted to the active agent in the
PT presence of the caspase -
XX
PS Disclosure; Page 13; 61pp; English.
XX
CC The present sequence is that of the preferred peptide substrate
CC of granzyme-B. The invention provides methods for the localised
CC delivery of pharmaceutical agents by the administration of a
CC caspase conjugate that targets a cell type of interest and the

CC additional administration of a pro-agent, which is locally
CC converted by the caspase to an active agent. In preferred
CC embodiments the targeting component is an antibody, the drug is
CC preferably a cytotoxic or chemotherapeutic agent, and the
CC prodrug moiety comprises a caspase-cleavable peptide, especially
CC a group II caspase cleavable peptide (see AAB82926), which is a
CC very poor substrate for granzyme B and for proinflammatory
CC caspases. The methods allow treatment of diseases or disorders
CC characterised by the appearance or presence of a particular cell
CC type. Such cells include bacterially and virally infected cells
CC expressing cell surface epitopes characteristic of the infection,
CC neoplastic and malignant cells such as tumour cells, and cells
CC characterised by their presence or appearance in areas of
CC inflammation.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 22; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
:||
Db 1 IEP 3

RESULT 12
AAG62561
ID AAG62561 standard; peptide; 4 AA.
XX
AC AAG62561;
XX
DT 05-SEP-2001 (first entry)
XX
DE Cresyl violet substituted caspase inhibitor substrate SEQ ID NO: 16.
XX
KW Caspase inhibitor substrate; cresyl violet; fluorogenic substrate;
KW metastasis; cancer; protease.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "optionally modified by carbobenzoxy or acetyl"
XX
PN US6235493-B1.
XX
PD 22-MAY-2001.
XX
PF 05-AUG-1998; 98US-0130193.
XX
PR 06-AUG-1997; 97US-0055392.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Bissell ER, Smith RF;
XX
DR WPI; 2001-431695/46.
XX
PT Method for detection of an enzyme in vivo or in vitro in a cell using
PT an amino acid-substituted cresyl violet fluorogenic substrate -
XX
PS Claim 6; Column 32; 27pp; English.
XX
CC The present invention describes a method of detecting the presence of an
CC enzyme in living cells, involving contacting the cell with an amino acid
CC substituted cresyl violet fluorogenic substrate and quantifying the
CC fluorescence produced. This is particularly useful for detecting enzymes
CC such as proteases and caspases, which may be linked to cancer metastasis.
CC The present sequence is a caspase inhibitor substrate used in the
CC exemplification of the invention.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 22; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
:||
Db 1 IEP 3

RESULT 13
AAB29759
ID AAB29759 standard; peptide; 4 AA.
XX
AC AAB29759;
XX
DT 28-FEB-2001 (first entry)
XX
DE Escherichia coli ribosomal protein L16 N-terminus.
XX
KW Recombinant human haemoglobin; rHb1.1; bacterial expression;
KW N-terminal methionine methylation signal; demethylation;
KW Escherichia coli; non-immunogenic; pharmaceutical composition.
XX
OS Escherichia coli.
XX
PN US6140071-A.
XX
PD 31-OCT-2000.
XX
PF 27-JAN-1994; 94US-0188374.
XX
PR 27-JAN-1994; 94US-0188374.
XX
PA (SOMA-) SOMATOGEN INC.
XX
PI Aitken JF, Apostol IZ, Levine JD, Lippincott JA;
XX
DR WPI; 2001-048957/06.
XX
PT Decreasing methylation of an N-terminus protein, especially hemoglobin
PT having proline at amino acid position 4, useful for producing
PT demethylated proteins for treating diseases, by altering this amino
PT acid to a non-proline residue -
XX
PS Disclosure; Column 1; 26pp; English.
XX
CC The invention relates to a method of decreasing the amount of N-terminal
CC methionine methylation on a protein expressed in a bacterium. The
CC bacterial methyltransferase which directs N-terminal methionine
CC methylation recognises proteins which have a proline residue at position
CC 4 (e.g., the bacterial ribosomal protein L16 and the bacterial chemotaxis
CC protein CheZ). The method comprises introducing mutations into the DNA
CC encoding the protein so that residue 4 is a non-proline residue, thereby
CC reducing the degree of N-terminal methylation when the protein is
CC expressed in a bacterium. The method is useful for decreasing methylation
CC of a protein, particularly a recombinantly produced protein. The
CC demethylated protein can be used in a pharmaceutical composition for
CC the treatment of a disease but with less likelihood of eliciting an
CC immunological response. These demethylated proteins may be used as
CC therapeutic agent for the treatment and/or amelioration of disease or
CC symptoms associated with a disease. The exemplifications describe
CC the expression of a recombinant human haemoglobin construct (rHb1.1) in
CC Escherichia coli, and its modification such that residue 4 of the
CC di-alpha chain of the recombinant haemoglobin is altered from proline
CC to a non-proline residue. The present sequence represents a peptide
CC referred to in the disclosure of the invention.
XX
SQ Sequence 4 AA;

Query Match 70.0%; Score 14; DB 22; Length 4;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;

```
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              :|:|
Db          1 MLQP 4


RESULT 14
AAE28993
ID   AAE28993 standard; peptide; 4 AA.
XX
AC   AAE28993;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Human thrombin cleavage site.
XX
KW   Vector; prokaryotic cell; human; thrombin cleavage site.
XX
OS   Homo sapiens.
XX
PN   WO200272847-A2.
XX
PD   19-SEP-2002.
XX
PF   22-FEB-2002; 2002WO-US05069.
XX
PR   09-MAR-2001; 2001US-274384P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Paegle ES,  Reilly D,  Yansura DG;
XX
DR   WPI; 2002-723363/78.
XX
PT   New vector comprising anti-termination nucleic acid or RNA encoding the
PT   polypeptide with a non-lambda promoter, useful for producing human
PT   thrombopoietin (TPO) or fibroblast growth factor-5 (FGF-5) polypeptide
PT   -
XX
PS   Example 1; Page 21; 70pp; English.
XX
CC   The invention relates to vectors for producing a polypeptide heterologous
CC   to prokaryotic cells and method for producing the polypeptide. The method
CC   is useful for producing a polypeptide heterologous to prokaryotic cells.
CC   The present sequence is human thrombin cleavage site used in the
CC   exemplification of the invention.
XX
SQ   Sequence   4 AA;

  Query Match             70.0%;  Score 14;  DB 23;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 15
ABG94471
ID   ABG94471 standard; Peptide; 4 AA.
XX
AC   ABG94471;
XX
DT   27-NOV-2002  (first entry)
XX
DE   Protease biosensor recognition site #24.
XX
KW   Detection; classification; identification; toxin detection;
KW   protease; ADP-ribosylating toxin; cytotoxic phospholipase;
KW   exfoliative toxin; toxic threat agent.
XX
OS   Synthetic.
XX
PN   US6416959-B1.
XX
PD   09-JUL-2002.
XX
PF   25-FEB-2000; 2000US-0513783.
XX
PR   26-FEB-1999;   99US-122152P.
PR   08-MAR-1999;   99US-123399P.
PR   31-AUG-1999;   99US-151797P.
PR   01-DEC-1999;   99US-168408P.
PR   27-FEB-1997;   97US-0810983.
PR   27-FEB-1998;   98US-0031271.
PR   12-JUL-1999;   99US-0352171.
PR   17-SEP-1999;   99US-0398965.
PR   29-OCT-1999;   99US-0430656.
XX
PA   (GIUL/) GIULIANO K.
PA   (KAPU/) KAPUR R.
XX
PI   Giuliano K,  Kapur R;
XX
DR   WPI; 2002-634730/68.
DR   N-PSDB; ABS71518.
XX
PT   Automated cell-based toxin detection, classification, and/or
PT   identification by treating cells involves use of three classes of
PT   luminescent reporter molecules such as detectors, classifiers or
PT   identifiers -
XX
PS   Example 10; Fig 29B; 214pp; English.
XX
CC   The invention describes methods of automated detection, classification
CC   and identification comprising treating cells containing luminescent
CC   reporter molecules (I) in array of locations with a test substance, where
CC   (I) are detectors, classifiers or identifiers, imaging cells in each
CC   location to obtain luminescent signals and converting optical information
CC   into digital data to interpret presence of toxins in the test substance.
CC   The method are useful for detection of toxins chosen from proteases,
CC   ADP-ribosylating toxins, cytotoxic phospholipases, and exfoliative
CC   toxins. Three classes of cell-based luminescent reporter molecules
CC   such as detectors, classifiers and identifiers are described and serve
CC   as reporters of toxic threat agents. The first two levels of
CC   characterisation ensure a rapid readout of toxin class without
CC   sacrificing the ability to detect many new mutant toxins or dissect
CC   several complex mixtures of known toxins. This is the amino acid sequence
CC   of a protease biosensor recognition site used in the cell-based screening
CC   system.
XX
SQ   Sequence   4 AA;

  Query Match             70.0%;  Score 14;  DB 23;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


Search completed: November 25, 2003, 14:06:39
Job time : 42 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:07:53 ; Search time 29 Seconds
(without alignments)
25.440 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 4771

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	4	9	US-09-729-873-16	Sequence 16, Appl
2	20	100.0	4	9	US-09-732-411-15	Sequence 15, Appl
3	20	100.0	4	15	US-10-091-724-7	Sequence 7, Appli
4	14	70.0	4	9	US-09-729-873-24	Sequence 24, Appl
5	14	70.0	4	10	US-09-947-387-23	Sequence 23, Appl
6	14	70.0	4	12	US-10-138-375-23	Sequence 23, Appl
7	14	70.0	4	15	US-10-100-957A-100	Sequence 100, App
8	14	70.0	4	15	US-10-080-866-6	Sequence 6, Appli
9	13	65.0	4	10	US-09-947-387-27	Sequence 27, Appl
10	13	65.0	4	12	US-09-932-165-1436	Sequence 1436, Ap
11	13	65.0	4	12	US-10-319-592-5	Sequence 5, Appli
12	13	65.0	4	12	US-10-138-375-27	Sequence 27, Appl
13	13	65.0	4	15	US-10-006-869-199	Sequence 199, App
14	12	60.0	3	12	US-09-834-240-32	Sequence 32, Appl
15	12	60.0	3	15	US-10-001-073-12	Sequence 12, Appl
16	12	60.0	4	10	US-09-947-387-13	Sequence 13, Appl
17	12	60.0	4	11	US-09-809-638-733	Sequence 733, App
18	12	60.0	4	11	US-09-852-910-169	Sequence 169, App
19	12	60.0	4	12	US-10-285-045-56	Sequence 56, Appl
20	12	60.0	4	12	US-10-076-047A-115	Sequence 115, App
21	12	60.0	4	12	US-10-284-660-56	Sequence 56, Appl
22	12	60.0	4	12	US-10-087-942-8	Sequence 8, Appli
23	12	60.0	4	12	US-10-196-394-120	Sequence 120, App
24	12	60.0	4	12	US-10-138-375-13	Sequence 13, Appl
25	12	60.0	4	15	US-10-087-905-8	Sequence 8, Appli
26	12	60.0	4	15	US-10-059-261-25	Sequence 25, Appl
27	12	60.0	4	15	US-10-059-261-133	Sequence 133, App
28	12	60.0	4	15	US-10-006-869-197	Sequence 197, App
29	12	60.0	4	15	US-10-205-270-4	Sequence 4, Appli
30	12	60.0	4	15	US-10-096-986-7	Sequence 7, Appli
31	12	60.0	4	15	US-10-096-986-31	Sequence 31, Appl
32	11	55.0	4	10	US-09-861-097-16	Sequence 16, Appl
33	11	55.0	4	11	US-09-861-098-16	Sequence 16, Appl
34	11	55.0	4	11	US-09-791-153A-42	Sequence 42, Appl
35	11	55.0	4	12	US-10-169-351-94	Sequence 94, Appl
36	11	55.0	4	12	US-10-022-066-163	Sequence 163, App
37	11	55.0	4	12	US-09-861-012-16	Sequence 16, Appl
38	11	55.0	4	12	US-10-277-292-677	Sequence 677, App
39	11	55.0	4	12	US-10-280-340-677	Sequence 677, App
40	11	55.0	4	15	US-10-006-869-269	Sequence 269, App
41	11	55.0	4	15	US-10-006-869-462	Sequence 462, App
42	11	55.0	4	15	US-10-006-869-476	Sequence 476, App
43	11	55.0	4	15	US-10-006-869-574	Sequence 574, App
44	11	55.0	4	15	US-10-012-456A-45	Sequence 45, Appl
45	10	50.0	3	10	US-09-031-629A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-729-873-16
; Sequence 16, Application US/09729873
; Patent No. US20010036921A1
; GENERAL INFORMATION:
; APPLICANT: Samy Ashkar
; TITLE OF INVENTION: Osteopontin-Derived Chemotactic and Inhibitory Agents
; TITLE OF INVENTION: and Uses Therefor
; FILE REFERENCE: CMZ-123CP
; CURRENT APPLICATION NUMBER: US/09/729,873
; CURRENT FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 60/129,764
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: PCT/US00/10344
; PRIOR FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: Peptide
US-09-729-873-16

Query Match 100.0%; Score 20; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEP 4
||||
Db 1 VLEP 4

RESULT 2
US-09-732-411-15
; Sequence 15, Application US/09732411

```
; Patent No. US20020058336A1
; GENERAL INFORMATION:
;  APPLICANT:  Samy Ashkar
;  TITLE OF INVENTION:  Adhesion Modulatory Peptides and Methods for Use
;  FILE REFERENCE:  CMZ-124CP
;  CURRENT APPLICATION NUMBER: US/09/732,411
;  CURRENT FILING DATE:  2000-12-07
;  PRIOR APPLICATION NUMBER:  60/129,709
;  PRIOR FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  PCT/US00/10329
;  PRIOR FILING DATE:  2000-04-17
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: peptide
US-09-732-411-15

  Query Match             100.0%;  Score 20;  DB 9;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 6e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          1 VLEP 4


RESULT 3
US-10-091-724-7
; Sequence 7, Application US/10091724
; Publication No. US20030105310A1
; GENERAL INFORMATION:
;  APPLICANT: Children's Medical Center Corporation
;  APPLICANT:  Ashkar, Samy
;  TITLE OF INVENTION: Method to Screen Peptide Libraries Using Minicell Display
;  FILE REFERENCE: CMCC 820
;  CURRENT APPLICATION NUMBER: US/10/091,724
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 60/306,946
;  PRIOR FILING DATE: 2001-07-20
;  PRIOR APPLICATION NUMBER: 60/274,039
;  PRIOR FILING DATE: 2001-03-07
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Inhibitor of macrophage recruitment.
US-10-091-724-7

  Query Match             100.0%;  Score 20;  DB 15;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 6e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              ||||
Db          1 VLEP 4


RESULT 4
US-09-729-873-24
; Sequence 24, Application US/09729873
; Patent No. US20010036921A1
; GENERAL INFORMATION:
;  APPLICANT:  Samy Ashkar
;  TITLE OF INVENTION:  Osteopontin-Derived Chemotactic and Inhibitory Agents
;  TITLE OF INVENTION:  and Uses Therefor
;  FILE REFERENCE:  CMZ-123CP
;  CURRENT APPLICATION NUMBER: US/09/729,873
;  CURRENT FILING DATE:  2000-12-05
;  PRIOR APPLICATION NUMBER:  60/129,764
;  PRIOR FILING DATE:  1999-04-15
;  PRIOR APPLICATION NUMBER:  PCT/US00/10344
;  PRIOR FILING DATE:  2000-04-17
;  NUMBER OF SEQ ID NOS:  25
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: Peptide
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: ACETYLATION
US-09-729-873-24

  Query Match             70.0%;  Score 14;  DB 9;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 6e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              :|:|
Db          1 MLDP 4


RESULT 5
US-09-947-387-23
; Sequence 23, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eckard
;  APPLICANT:  Cai, Sui Xiong
;  APPLICANT:  Keana, John F.W.
;  APPLICANT:  Drewe, John A.
;  APPLICANT:  Zhang, Han-Zhong
;  TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Molecu
;  TITLE OF INVENTION:  Their Applications for Whole-Cell Fluorescence
;  TITLE OF INVENTION:  Screening Assays for Caspases and Other Enzymes and the
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 1735.0290005
;  CURRENT APPLICATION NUMBER: US/09/947,387
;  CURRENT FILING DATE:  2001-09-07
;  PRIOR APPLICATION NUMBER: US 60/061,582
;  PRIOR FILING DATE: 1997-10-10
;  PRIOR APPLICATION NUMBER: US 60/145,746
;  PRIOR FILING DATE: 1998-03-03
;  PRIOR APPLICATION NUMBER: US 09/168,888
;  PRIOR FILING DATE: 1998-10-09
;  NUMBER OF SEQ ID NOS: 142
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-947-387-23

  Query Match             70.0%;  Score 14;  DB 10;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3
```

```
RESULT 6
US-10-138-375-23
; Sequence 23, Application US/10138375
; Publication No. US20030208037A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Han-Zhong
;  APPLICANT:  Cai, Sui Xiong
;  APPLICANT:  Drewe, John A.
;  APPLICANT:  Yang, Wu
;  TITLE OF INVENTION: No. US20030208037A1el Fluorescence Dyes and Their Applications fo
;  TITLE OF INVENTION:  Fluorescence Screening Assays for Caspases, Peptidases, Protease
;  TITLE OF INVENTION:  Other Enzymes and the Use Thereof
;  FILE REFERENCE: 1735.0030001
;  CURRENT APPLICATION NUMBER: US/10/138,375
;  CURRENT FILING DATE:  2002-05-06
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/357,952
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/093,642
;  PRIOR FILING DATE: EARLIER FILING DATE: 21-JUL-1998
;  NUMBER OF SEQ ID NOS: 139
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: Peptide
US-10-138-375-23

  Query Match             70.0%;  Score 14;  DB 12;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 7
US-10-100-957A-100
; Sequence 100, Application US/10100957A
; Publication No. US20030096322A1
; GENERAL INFORMATION:
;  APPLICANT: Giuliano, Kenneth A.
;  APPLICANT:  Kapur, Ravi
;  TITLE OF INVENTION: A System for Cell Based Screening
;  FILE REFERENCE: 97-022-L1A
;  CURRENT APPLICATION NUMBER: US/10/100,957A
;  CURRENT FILING DATE:  2002-03-19
;  NUMBER OF SEQ ID NOS: 180
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Granzyme B
;   OTHER INFORMATION: substrate recognition sequence
US-10-100-957A-100

  Query Match             70.0%;  Score 14;  DB 15;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 8
US-10-080-866-6
; Sequence 6, Application US/10080866
; Publication No. US20030109024A1
; GENERAL INFORMATION:
;  APPLICANT: Paegle, E. Sasha
;  APPLICANT:  Reilly, Dorothea
;  APPLICANT:  Yansura, Daniel G.
;  TITLE OF INVENTION: PROCESS FOR PRODUCTION OF POLYPEPTIDES
;  FILE REFERENCE: P1732R1
;  CURRENT APPLICATION NUMBER: US/10/080,866
;  CURRENT FILING DATE:  2002-02-22
;  PRIOR APPLICATION NUMBER: US 60/274,384
;  PRIOR FILING DATE: 2001-03-09
;  NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 6
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-080-866-6

  Query Match             70.0%;  Score 14;  DB 15;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 9
US-09-947-387-27
; Sequence 27, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eckard
;  APPLICANT:  Cai, Sui Xiong
;  APPLICANT:  Keana, John F.W.
;  APPLICANT:  Drewe, John A.
;  APPLICANT:  Zhang, Han-Zhong
;  TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Molecul
;  TITLE OF INVENTION:  Their Applications for Whole-Cell Fluorescence
;  TITLE OF INVENTION:  Screening Assays for Caspases and Other Enzymes and the
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 1735.0290005
;  CURRENT APPLICATION NUMBER: US/09/947,387
;  CURRENT FILING DATE:  2001-09-07
;  PRIOR APPLICATION NUMBER: US 60/061,582
;  PRIOR FILING DATE: 1997-10-10
;  PRIOR APPLICATION NUMBER: US 60/145,746
;  PRIOR FILING DATE: 1998-03-03
;  PRIOR APPLICATION NUMBER: US 09/168,888
;  PRIOR FILING DATE: 1998-10-09
;  NUMBER OF SEQ ID NOS: 142
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-947-387-27

  Query Match             65.0%;  Score 13;  DB 10;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 VEP 3


RESULT 10
US-09-932-165-1436
; Sequence 1436, Application US/09932165
```

; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALLITA-EID, PIA M.
; APPLICANT: FARIS, MARY
; APPLICANT: SAFFRAN, DOUGLAS
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITS, AYA
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CaTrF2E11 USEFUL IN TREATMENT AND
; TITLE OF INVENTION: DETECTION OF CANCER
; FILE REFERENCE: 51158-20014.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1436
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-09-932-165-1436

Query Match 65.0%; Score 13; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLE 3
|||
Db 2 VLE 4

RESULT 11
US-10-319-592-5
; Sequence 5, Application US/10319592
; Publication No. US20030152567A1
; GENERAL INFORMATION:
; APPLICANT: ROMISCH, JURGEN
; APPLICANT: FEUBNER, ANNETTE
; APPLICANT: STOHR, HANS-ARNOLD
; TITLE OF INVENTION: PROTEASE FOR ACTIVATING CLOTTING FACTOR VII
; FILE REFERENCE: 06478.1424
; CURRENT APPLICATION NUMBER: US/10/319,592
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US/09/295,316
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 19818495.6
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 19827734.2
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 19851332.1
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 19851336.4
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 19851335.6
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 19903693.4
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Illustrative
; OTHER INFORMATION: peptide
US-10-319-592-5

Query Match 65.0%; Score 13; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
|:|
Db 2 LDP 4

RESULT 12
US-10-138-375-27
; Sequence 27, Application US/10138375
; Publication No. US20030208037A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Han-Zhong
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: No. US20030208037A1el Fluorescence Dyes and Their Applications f
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Proteas
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.0030001
; CURRENT APPLICATION NUMBER: US/10/138,375
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/357,952
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/093,642
; PRIOR FILING DATE: EARLIER FILING DATE: 21-JUL-1998
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: Peptide
US-10-138-375-27

Query Match 65.0%; Score 13; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEP 4
:||
Db 1 VEP 3

RESULT 13
US-10-006-869-199
; Sequence 199, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative linear modulating agent based on
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-10-006-869-199

Query Match 65.0%; Score 13; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 LEP 4
              :||
Db          1 VEP 3
```

RESULT 14
US-09-834-240-32
; Sequence 32, Application US/09834240
; Publication No. US20030166261A1
; GENERAL INFORMATION:
; APPLICANT: Sompuram, Seshi R.
; APPLICANT: Ramanathan, Halasya
; TITLE OF INVENTION: Quality Control for Cytochemical Assays
; FILE REFERENCE: 1159.1008-005
; CURRENT APPLICATION NUMBER: US/09/834,240
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 09/549,855
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 09/291,351
; PRIOR FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-240-32

Query Match 60.0%; Score 12; DB 12; Length 3;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 EP 4
              ||
Db          2 EP 3
```

RESULT 15
US-10-001-073-12
; Sequence 12, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; APPLICANT: Small, Kirsten
; TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
; FILE REFERENCE: 13073-PCT
; CURRENT APPLICATION NUMBER: US/10/001,073
; CURRENT FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-001-073-12

Query Match 60.0%; Score 12; DB 15; Length 3;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 EP 4
              ||
Db          2 EP 3
```

Search completed: November 25, 2003, 14:13:04
Job time : 30 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:05:52 ; Search time 21 Seconds
(without alignments)
8.059 Million cell updates/sec

Title: US-09-732-411-15
Perfect score: 20
Sequence: 1 VLEP 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 10151

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15	75.0	4	1	US-08-240-712-5	Sequence 5, Appli
2	15	75.0	4	1	US-08-443-890-5	Sequence 5, Appli
3	15	75.0	4	3	US-09-058-562-5	Sequence 5, Appli
4	15	75.0	4	5	PCT-US92-09752-5	Sequence 5, Appli
5	14	70.0	4	1	US-08-336-343A-38	Sequence 38, Appl
6	14	70.0	4	2	US-08-609-271-1	Sequence 1, Appli
7	14	70.0	4	3	US-08-188-374-1	Sequence 1, Appli
8	14	70.0	4	3	US-09-130-193-16	Sequence 16, Appl
9	14	70.0	4	3	US-09-357-952-23	Sequence 23, Appl
10	14	70.0	4	4	US-09-521-650-23	Sequence 23, Appl
11	14	70.0	4	4	US-09-168-888-23	Sequence 23, Appl
12	14	70.0	4	4	US-09-513-783A-100	Sequence 100, App
13	13	65.0	4	2	US-08-441-871-28	Sequence 28, Appl
14	13	65.0	4	3	US-09-330-970-16	Sequence 16, Appl
15	13	65.0	4	3	US-09-330-970-34	Sequence 34, Appl
16	13	65.0	4	3	US-08-859-242-31	Sequence 31, Appl
17	13	65.0	4	3	US-09-357-952-27	Sequence 27, Appl
18	13	65.0	4	4	US-09-521-650-27	Sequence 27, Appl
19	13	65.0	4	4	US-09-168-888-27	Sequence 27, Appl
20	13	65.0	4	4	US-09-187-859-199	Sequence 199, App
21	13	65.0	4	4	US-09-295-316-5	Sequence 5, Appli
22	13	65.0	4	4	US-09-446-787B-45	Sequence 45, Appl
23	13	65.0	4	4	US-09-839-542B-199	Sequence 199, App
24	12	60.0	4	1	US-07-714-167E-8	Sequence 8, Appli
25	12	60.0	4	1	US-07-714-167E-11	Sequence 11, Appl
26	12	60.0	4	1	US-07-729-353-7	Sequence 7, Appli
27	12	60.0	4	1	US-08-122-546-1	Sequence 1, Appli
28	12	60.0	4	1	US-08-280-443-39	Sequence 39, Appl
29	12	60.0	4	1	US-08-457-459-39	Sequence 39, Appl
30	12	60.0	4	1	US-08-224-868-5	Sequence 5, Appli
31	12	60.0	4	1	US-08-555-678-39	Sequence 39, Appl
32	12	60.0	4	1	US-08-477-509B-7	Sequence 7, Appli
33	12	60.0	4	1	US-08-477-509B-31	Sequence 31, Appl
34	12	60.0	4	2	US-08-609-271-2	Sequence 2, Appli
35	12	60.0	4	2	US-08-609-271-3	Sequence 3, Appli
36	12	60.0	4	2	US-08-609-271-4	Sequence 4, Appli
37	12	60.0	4	2	US-08-609-271-5	Sequence 5, Appli
38	12	60.0	4	2	US-08-764-938-1	Sequence 1, Appli
39	12	60.0	4	3	US-08-482-085B-7	Sequence 7, Appli
40	12	60.0	4	3	US-08-482-085B-31	Sequence 31, Appl
41	12	60.0	4	3	US-08-415-655-3	Sequence 3, Appli
42	12	60.0	4	3	US-08-624-405-7	Sequence 7, Appli
43	12	60.0	4	3	US-09-105-678A-50	Sequence 50, Appl
44	12	60.0	4	3	US-09-131-052-1	Sequence 1, Appli
45	12	60.0	4	3	US-08-188-374-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-240-712-5
; Sequence 5, Application US/08240712
; Patent No. 5599907
;  GENERAL INFORMATION:
;    APPLICANT:  ANDERSON, DAVID C.
;    APPLICANT:  MATHEWS, ANTONY JAMES
;    APPLICANT:  STETLER, GARY L.
;    TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC
;    TITLE OF INVENTION:  HEMOGLOBINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/240,712
;      FILING DATE:  09-MAY-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/09752
;      FILING DATE:  13-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANDERSON=6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-240-712-5

  Query Match           75.0%;  Score 15;  DB 1;  Length 4;
  Best Local Similarity  75.0%;  Pred. No. 2.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 VLEP 4
              || |
Db          1 VLSP 4


RESULT 2
US-08-443-890-5
; Sequence 5, Application US/08443890
; Patent No. 5739011
;  GENERAL INFORMATION:
;    APPLICANT:  ANDERSON, DAVID C.
;    APPLICANT:  MATHEWS, ANTONY JAMES
;    APPLICANT:  STETLER, GARY L.
;    TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC
;    TITLE OF INVENTION:  HEMOGLOBINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/443,890
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/240,712
;      FILING DATE:  09-MAY-1994
;      APPLICATION NUMBER:  PCT/US92/09752
;      FILING DATE:  13-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANDERSON=6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-443-890-5

  Query Match             75.0%;  Score 15;  DB 1;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              || |
Db          1 VLSP 4


RESULT 3
US-09-058-562-5
; Sequence 5, Application US/09058562A
; Patent No. 6184356
; GENERAL INFORMATION:
;  APPLICANT:  Anderson, David C.
;  APPLICANT:  Mathews, Antony James
;  APPLICANT:  Stetler, Gary L.
;  TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC HEMOGLOBINS
;  FILE REFERENCE:  BXTB 2087
;  CURRENT APPLICATION NUMBER: US/09/058,562A
;  CURRENT FILING DATE:  1998-04-13
;  PRIOR APPLICATION NUMBER:  US 08/240,712
;  PRIOR FILING DATE:  1994-05-09
;  PRIOR APPLICATION NUMBER:  PCT/US92/09752
;  PRIOR FILING DATE:  1993-05-13
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: peptide
US-09-058-562-5

  Query Match             75.0%;  Score 15;  DB 3;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              || |
Db          1 VLSP 4


RESULT 4
PCT-US92-09752-5
; Sequence 5, Application PC/TUS9209752
;  GENERAL INFORMATION:
;    APPLICANT:  ANDERSON, DAVID C.
;    APPLICANT:  MATHEWS, ANTONY JAMES
;    APPLICANT:  STETLER, GARY L.
;    TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC
;    TITLE OF INVENTION:  HEMOGLOBINS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/09752
;      FILING DATE:  19930109
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANDERSON-6-PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US92-09752-5

  Query Match             75.0%;  Score 15;  DB 5;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 VLEP 4
              || |
Db          1 VLSP 4


RESULT 5
US-08-336-343A-38
; Sequence 38, Application US/08336343A
; Patent No. 5677144
;  GENERAL INFORMATION:
;    APPLICANT:  Ullrich, Axel
;    APPLICANT:  Alves, Frauke
;    TITLE OF INVENTION:  CCK-2, A No. 5677144el Receptor Tyrosine Kinase
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/336,343A
;      FILING DATE:  08-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7683-065
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-336-343A-38

  Query Match             70.0%;  Score 14;  DB 1;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          2 MEP 4


RESULT 6
US-08-609-271-1
; Sequence 1, Application US/08609271
; Patent No. 5811264
;  GENERAL INFORMATION:
;    APPLICANT:  Aitken, Jacqueline F.
;    APPLICANT:  Apostol, Izydor Z.
;    APPLICANT:  Lippincott, Julie A.
;    APPLICANT:  Levine, Joseph D.
;    TITLE OF INVENTION:  Proteins with Mutations to Decrease N-Terminal Methylation
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Somatogen, Inc.
;      STREET:  2545 Central Avenue, Site FD-1
;      CITY:  Boulder
;      STATE:  Colorado
;      ZIP:  80301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.0.1
;      SOFTWARE:  Microsoft Word 5.0a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/609,271
;      FILING DATE:  28 February 1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/188,374
;      FILING DATE:  1/27/94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5811264elli, Marianne F.
;      REGISTRATION NUMBER:  38571
;      NAME:  Brown,Theresa A.
;      REGISTRATION NUMBER:  32547
;      REFERENCE/DOCKET NUMBER:  170/Div
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303-541-3324
;      TELEFAX:  303-444-3013
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  no
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label=Modified-site1
;      OTHER INFORMATION:  /note= "Met(1) is Methylated N terminal Met"
US-08-609-271-1

  Query Match             70.0%;  Score 14;  DB 2;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLEP 4
              :|:|
Db          1 MLQP 4


RESULT 7
US-08-188-374-1
; Sequence 1, Application US/08188374B
; Patent No. 6140071
; GENERAL INFORMATION:
;  APPLICANT: Aitken F., Jacqueline
;  APPLICANT: Apostol, Izydor Z.
;  APPLICANT: Lippincott, Julie A.
;  APPLICANT: Levine, Joseph D.
;  TITLE OF INVENTION: Proteins with Mutations to Decrease N-Terminal Meth
;  FILE REFERENCE: BXTB 1953
;  CURRENT APPLICATION NUMBER: US/08/188,374B
;  CURRENT FILING DATE: 1994-01-27
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: METHYLATION- N terminal Met
US-08-188-374-1

  Query Match             70.0%;  Score 14;  DB 3;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 VLEP 4
              :|:|
Db          1 MLQP 4


RESULT 8
US-09-130-193-16
; Sequence 16, Application US/09130193
; Patent No. 6235493
;  GENERAL INFORMATION:
;    APPLICANT:  EUGENE R. BISSELL ET AL
;    TITLE OF INVENTION:  AMINO ACID SUBSTITUTED-
;    TITLE OF INVENTION:  CRESYL VIOLET, SYNTHETIC
;    TITLE OF INVENTION:  FLUOROGENIC SUBSTRATES
;    TITLE OF INVENTION:  FOR THE ANALYSIS OF
;    TITLE OF INVENTION:  AGENTS IN INDIVIDUAL IN
;    TITLE OF INVENTION:  VIVO CELLS OR TISSUE
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIK A, LLP
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,193
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/026,062
;      FILING DATE:  SEPTEMBER 13, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOWARD m. PETERS
;      REGISTRATION NUMBER:  29,202
;      REFERENCE/DOCKET NUMBER:  3586.01-1(HMP)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-1677 X 20
;      TELEFAX:  (650) 324-1678
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-193-16

  Query Match             70.0%;  Score 14;  DB 3;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 9
US-09-357-952-23
; Sequence 23, Application US/09357952
; Patent No. 6248904
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Han-Zhong
;  APPLICANT: Cai, Sui Xiong
;  APPLICANT: Drewe, John A.
;  APPLICANT: Yang, Wu
;  TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for Whole-
;  TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Proteases
;  TITLE OF INVENTION: Other Enzymes and the Use Thereof
;  FILE REFERENCE: 1735.0030001
;  CURRENT APPLICATION NUMBER: US/09/357,952
;  CURRENT FILING DATE: 1999-07-21
;  EARLIER APPLICATION NUMBER: US 60/093,642
;  EARLIER FILING DATE: 21-JUL-1998
;  NUMBER OF SEQ ID NOS: 139
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: Peptide
US-09-357-952-23

  Query Match             70.0%;  Score 14;  DB 3;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 10
US-09-521-650-23
; Sequence 23, Application US/09521650
; Patent No. 6335429
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eckard
;  APPLICANT: Cai, Sui Xiong
;  APPLICANT: Keana, John F.W.
;  APPLICANT: Drewe, John A.
;  APPLICANT: Zhang, Han-Zhong
;  TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and
;  TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
;  TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
;  TITLE OF INVENTION: Use Thereof
;  FILE REFERENCE: 1735.0290002
;  CURRENT APPLICATION NUMBER: US/09/521,650
;  CURRENT FILING DATE: 2000-03-08
;  EARLIER APPLICATION NUMBER: 09/168,888
;  EARLIER FILING DATE: 1998-10-09
;  EARLIER APPLICATION NUMBER: US 60/061,582
;  EARLIER FILING DATE: 1997-10-10
;  EARLIER APPLICATION NUMBER: US 09/033,661
;  EARLIER FILING DATE: 1998-03-03
;  NUMBER OF SEQ ID NOS: 142
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: Peptide
US-09-521-650-23

  Query Match             70.0%;  Score 14;  DB 4;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 11
US-09-168-888-23
; Sequence 23, Application US/09168888
; Patent No. 6342611
```

```
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eckard
;  APPLICANT: Cai, Sui Xiong
;  APPLICANT: Keana, John F.W.
;  APPLICANT: Drewe, John A.
;  APPLICANT: Zhang, Han-Zhong
;  TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and
;  TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
;  TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
;  TITLE OF INVENTION: Use Thereof
;  FILE REFERENCE: 1735.0290002
;  CURRENT APPLICATION NUMBER: US/09/168,888
;  CURRENT FILING DATE: 1998-10-09
;  EARLIER APPLICATION NUMBER: US 60/061,582
;  EARLIER FILING DATE: 1997-10-10
;  EARLIER APPLICATION NUMBER: US 09/033,661
;  EARLIER FILING DATE: 1998-03-03
;  NUMBER OF SEQ ID NOS: 142
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: Peptide
US-09-168-888-23

  Query Match             70.0%;  Score 14;  DB 4;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 12
US-09-513-783A-100
; Sequence 100, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
;  APPLICANT:  Giuliano, Kenneth A.
;  APPLICANT:  Kapur, Ravi
;  TITLE OF INVENTION:  A System for Cell Based Screening
;  FILE REFERENCE:  97-022-L1
;  CURRENT APPLICATION NUMBER: US/09/513,783A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS:  180
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 100
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Granzyme B
;   OTHER INFORMATION: substrate recognition sequence
US-09-513-783A-100

  Query Match             70.0%;  Score 14;  DB 4;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              :||
Db          1 IEP 3


RESULT 13
US-08-441-871-28
; Sequence 28, Application US/08441871
; Patent No. 5846765
;  GENERAL INFORMATION:
;    APPLICANT:  Matthews, David J.
;    APPLICANT:  Wells, James A.
;    APPLICANT:  Zoller, Mark J.
;    TITLE OF INVENTION:  Identification of No. 5846765el Substrates
;    NUMBER OF SEQUENCES:  152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/441,871
;      FILING DATE:  16-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/418928
;      FILING DATE:  05-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/161692
;      FILING DATE:  03-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/864452
;      FILING DATE:  06-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/09133
;      FILING DATE:  03-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/743614
;      FILING DATE:  09-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/715300
;      FILING DATE:  14-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/683400
;      FILING DATE:  10-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/621667
;      FILING DATE:  03-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Daryl B.
;      REGISTRATION NUMBER:  32,637
;      REFERENCE/DOCKET NUMBER:  645P5C2D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1249
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-441-871-28

  Query Match             65.0%;  Score 13;  DB 2;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LEP 4
              |:|
Db          2 LDP 4


RESULT 14
US-09-330-970-16
; Sequence 16, Application US/09330970
; Patent No. 6146876
```

```
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  APPLICANT: Kapeller-Libermann, Rosana
;  APPLICANT: White, David
;  TITLE OF INVENTION: A No. 6146876el Human Cyclic Nucleotide
;  TITLE OF INVENTION: Phophodiesterase
;  FILE REFERENCE: 5800-28
;  CURRENT APPLICATION NUMBER: US/09/330,970
;  CURRENT FILING DATE: 1999-06-11
;  EARLIER APPLICATION NUMBER: 09/277,423
;  EARLIER FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-330-970-16

  Query Match             65.0%;  Score 13;  DB 3;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLE 3
              |||
Db          2 VLE 4


RESULT 15
US-09-330-970-34
; Sequence 34, Application US/09330970
; Patent No. 6146876
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  APPLICANT: Kapeller-Libermann, Rosana
;  APPLICANT: White, David
;  TITLE OF INVENTION: A No. 6146876el Human Cyclic Nucleotide
;  TITLE OF INVENTION: Phophodiesterase
;  FILE REFERENCE: 5800-28
;  CURRENT APPLICATION NUMBER: US/09/330,970
;  CURRENT FILING DATE: 1999-06-11
;  EARLIER APPLICATION NUMBER: 09/277,423
;  EARLIER FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 34
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-330-970-34

  Query Match             65.0%;  Score 13;  DB 3;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VLE 3
              |||
Db          2 VLE 4


Search completed: November 25, 2003, 14:08:57
Job time : 22 secs
```